



Aalborg Universitet

AALBORG UNIVERSITY
DENMARK

Complex spatio-temporal distribution and genomic ancestry of mitochondrial DNA haplogroups in 24,216 Danes

Bybjerg-Grauholm, Jonas; Hagen, Christian M; Gonçalves, Vanessa F; Bækvad-Hansen, Marie; Hansen, Christine S; Hedley, Paula L; Kanter, Jørgen K; Nielsen, Jimmi; Theisen, Michael; Mors, Ole; Kennedy, James; Als, Thomas D; Demur, Alfonso B; Nordentoft, Merete; Børghlum, Anders; Mortensen, Preben B; Werge, Thomas M; Hougaard, David M; Christiansen, Michael

Published in:
PLOS ONE

DOI (link to publication from Publisher):
[10.1371/journal.pone.0208829](https://doi.org/10.1371/journal.pone.0208829)

Creative Commons License
CC BY 4.0

Publication date:
2018

Document Version
Publisher's PDF, also known as Version of record

[Link to publication from Aalborg University](#)

Citation for published version (APA):

Bybjerg-Grauholm, J., Hagen, C. M., Gonçalves, V. F., Bækvad-Hansen, M., Hansen, C. S., Hedley, P. L., Kanter, J. K., Nielsen, J., Theisen, M., Mors, O., Kennedy, J., Als, T. D., Demur, A. B., Nordentoft, M., Børghlum, A., Mortensen, P. B., Werge, T. M., Hougaard, D. M., & Christiansen, M. (2018). Complex spatio-temporal distribution and genomic ancestry of mitochondrial DNA haplogroups in 24,216 Danes. *PLOS ONE*, 13(12), [e0208829]. <https://doi.org/10.1371/journal.pone.0208829>

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- ? Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- ? You may not further distribute the material or use it for any profit-making activity or commercial gain
- ? You may freely distribute the URL identifying the publication in the public portal ?

RESEARCH ARTICLE

Complex spatio-temporal distribution and genomic ancestry of mitochondrial DNA haplogroups in 24,216 Danes

Jonas Bybjerg-Grauholm¹, Christian M. Hagen¹, Vanessa F. Gonçalves², Marie Bækvad-Hansen¹, Christine S. Hansen¹, Paula L. Hedley¹, Jørgen K. Kanthers³, Jimmi Nielsen⁴, Michael Theisen¹, Ole Mors⁵, James Kennedy², Thomas D. Als⁶, Alfonso B. Demur⁷, Merete Nordentoft⁸, Anders Børghlum⁶, Preben B. Mortensen⁹, Thomas M. Werge⁷, David M. Hougaard¹, Michael Christiansen^{1,3*}

1 Department for Congenital Disorders, Statens Serum Institut, Copenhagen, Denmark, **2** Centre for Addiction and Mental Health, University of Toronto, Toronto, Canada, **3** Department of Biomedical Sciences, University of Copenhagen, Copenhagen, Denmark, **4** Aalborg Psychiatric Hospital, Aalborg University Hospital, Aalborg, Denmark, **5** Department of Clinical Medicine, Aarhus University, Aarhus, Denmark, **6** Institute of Medical Genetics, Aarhus University, Aarhus, Denmark, **7** Mental Health Centre, Sct Hans, Capital Region of Denmark, Denmark, **8** Mental Health Centre, Capital Region of Denmark, Denmark, **9** Center for Register Research, Institute of Economics, Aarhus University, Århus, Denmark

☞ These authors contributed equally to this work.

* mic@ssi.dk



OPEN ACCESS

Citation: Bybjerg-Grauholm J, Hagen CM, Gonçalves VF, Bækvad-Hansen M, Hansen CS, Hedley PL, et al. (2018) Complex spatio-temporal distribution and genomic ancestry of mitochondrial DNA haplogroups in 24,216 Danes. PLoS ONE 13 (12): e0208829. <https://doi.org/10.1371/journal.pone.0208829>

Editor: Dan E. Arking, Johns Hopkins University, UNITED STATES

Received: June 21, 2018

Accepted: November 23, 2018

Published: December 13, 2018

Copyright: © 2018 Bybjerg-Grauholm et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: Due to the nature of the participant consent used in data collection, data are restricted by the scientific ethics committee and the iPSYCH group (www.ipsych.au.dk). Data can be made available through contact with corresponding author or Annette B Rasmussen, iPSYCH (abr@biomed.au.dk).

Funding: The iPSYCH study was funded by The Lundbeck Foundation Initiative for Integrative Psychiatric Research. This research has been

Abstract

Mitochondrial DNA (mtDNA) haplogroups (hgs) are evolutionarily conserved sets of mtDNA SNP-haplotypes with characteristic geographical distribution. Associations of hgs with disease and physiological characteristics have been reported, but have frequently not been reproducible. Using 418 mtDNA SNPs on the PsychChip (Illumina), we assessed the spatio-temporal distribution of mtDNA hgs in Denmark from DNA isolated from 24,642 geographically un-biased dried blood spots (DBS), collected from 1981 to 2005 through the Danish National Neonatal Screening program. ADMIXTURE was used to establish the genomic ancestry of all samples using a reference of 100K+ autosomal SNPs in 2,248 individuals from nine populations. Median-joining analysis determined that the hgs were highly variable, despite being typically Northern European in origin, suggesting multiple founder events. Furthermore, considerable heterogeneity and variation in nuclear genomic ancestry was observed. Thus, individuals with hg H exhibited 95%, and U hgs 38.2% - 92.5%, Danish ancestry. Significant clines between geographical regions and rural and metropolitan populations were found. Over 25 years, macro-hg L increased from 0.2% to 1.2% ($p = 1.1 \times 10^{-10}$), and M from 1% to 2.4% ($p = 3.7 \times 10^{-8}$). Hg U increased among the R macro-hg from 14.1% to 16.5% ($p = 1.9 \times 10^{-3}$). Genomic ancestry, geographical skewedness, and sub-hg distribution suggested that the L, M and U increases are due to immigration. The complex spatio-temporal dynamics and genomic ancestry of mtDNA in the Danish population reflect repeated migratory events and, in later years, net immigration. Such complexity may explain the often contradictory and population-specific reports of mito-genomic association with disease.

conducted using the Danish National Biobank resource, supported by the Novo Nordisk Foundation. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: The authors have declared that no competing interests exist.

Introduction

Mitochondria are subcellular organelles responsible for oxidative phosphorylation (OXPHOS), producing ~ 80% of the ATP in eukaryotic cells[1], apoptosis and cell-cycle regulation[2], redox- and calcium homeostasis[3] as well as intracellular signaling[4]. Each mitochondrion contains 2–10 copies of a 16.6 kb double-stranded mtDNA containing 37 genes[5]. Thirteen genes code for proteins in the five enzyme complexes conducting OXPHOS, whereas twenty-two genes code for tRNAs and two for rRNAs, all involved in intra-mitochondrial translation[6]. The mitochondrial proteome comprises approximately 1200 proteins[7, 8], of which mtDNA genes encode ~ 1%. The mtDNA is maternally inherited[9], exhibits a high mutation rate[10], and does not undergo recombination[5]. Genetic variants in mtDNA—as well as variants in the nuclear genes encoding the mitochondrial proteome—have been associated with disease[11–13]. Over 150 mitochondrial syndromes[14] have been associated with over 300 variants[7, 15].

Geographically and population specific lineages of mtDNA, haplogroups (hgs), have become fixed[16], through the processes of random genetic drift and selection as the human populations dispersed throughout the world[17]. The advent of high throughput DNA sequencing technology, as well as implementation of biobanking technologies[18], has enabled the construction of a high-resolution phylogenetic matrilineal mtDNA tree (Fig 1)[19].

Mitochondrial hgs have been assigned a role as disease modifiers[5]. Particularly in neurological degenerative diseases such as Alzheimer's disease[20–23] and Parkinson's disease[23–25], but also in psychiatric disease[26] and cardiac diseases such as hypertrophic[27, 28] and ischemic cardiomyopathy[29]. Supporting a role as disease modifiers, some mtDNA hgs have specific physiological characteristics, e.g. reduced or increased ATP synthesis rates[30, 31], and variation in methylation status of genes involved in inflammation and signaling[32, 33].

Most countries have a complex history with repeated migrations[34] and several bottlenecks caused by disease, war and emigration[35, 36]. These demographic events are reflected in the fine scale genetic structure within sub-populations[37]. However, the significance of such events to countrywide mtDNA hg distribution has not yet been assessed. As mtDNA interact functionally with the nuclear genome, it is paramount to ensure that specific mtDNA hgs—which are a marker of matrilineal genetic origin—do not represent population sub-structure at the genomic level. In theory, mtDNA is inherited independently of the nuclear genome, but population admixture and geographic isolation may result in bi-genomic dependence between mtDNA and the nuclear genome which might interfere with genetic association analysis.

Here we demonstrate the complexity of the spatio-temporal dynamics and genomic ancestry of mtDNA hgs in 24,216 Danes, which were sampled at birth during a 25-year period. This number represents 1.6% of the population. DNA was extracted from dried blood spot (DBS) samples obtained as part of the Danish Neonatal Screening Program[38], the very nature of which makes sampling geographically un-biased (Fig 2). Array analysis was performed using the PsychChip (Illumina, CA, USA) typing 588,454 variants.

Results

Distribution of macro-haplogroups

The mtDNA macro-hg distribution pattern is typically northern European[39] with > 90% belonging to the R macro-hg, 7.0% belonging to N and 1.6% to M of likely Near Eastern or Asian origin[39] and 0.7% belonging to the combined L macro-hgs (L0–L6) of a likely African origin[40] (Table 1). The hg defining SNPs are given in S1 Table and the proportion of samples not haplotypable in S2 Table. A PCA based on all the mtDNA SNPs showed a clear separation

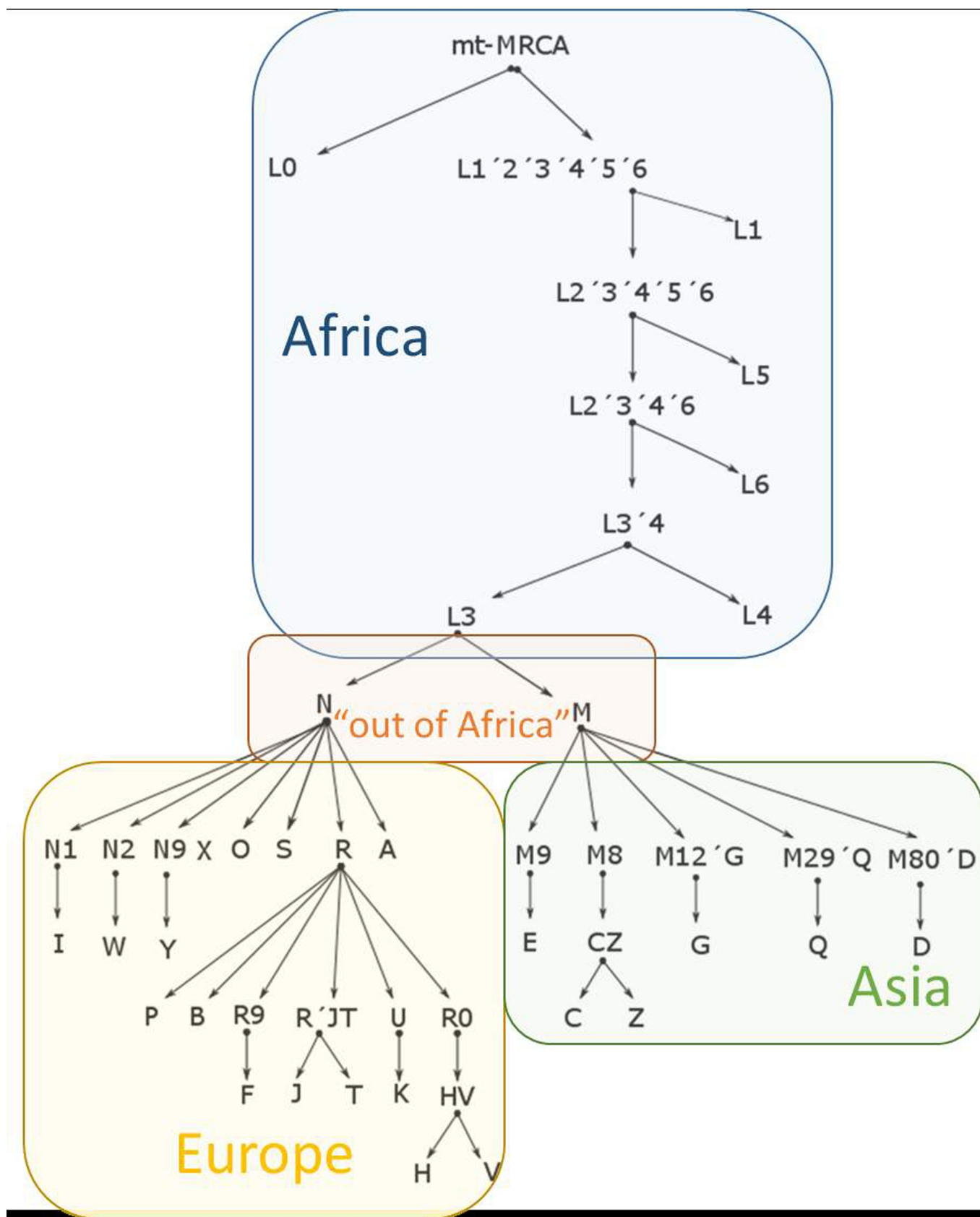


Fig 1. Phylogenetic tree of mtDNA sequences, modified from (<http://www.phylotree.org>). Approximate geographic origins of the haplogroups are represented by colored blocks. MRCA: Most recent common ancestor.

<https://doi.org/10.1371/journal.pone.0208829.g001>

between R and L, with N and M located intermediately (Fig 3). The PC1 seems to reflect time since branching, whereas PC2 reflects geographical distance.

Distribution of haplogroups

The R macro-hg was dissolved into hgs (Table 2) however, it was not possible, with the available SNPs, to differentiate the HV and P hgs from the R macro-hg (S1 Table, for details). Principal Component 1 versus 3 plot (Fig 4), demonstrated a clear clustering of mtDNA SNPs from persons belonging to each hg. The proximity of the clusters comprising U and K and H and V, respectively, is in accordance with the current phylogenetic mtDNA tree (Fig 1).

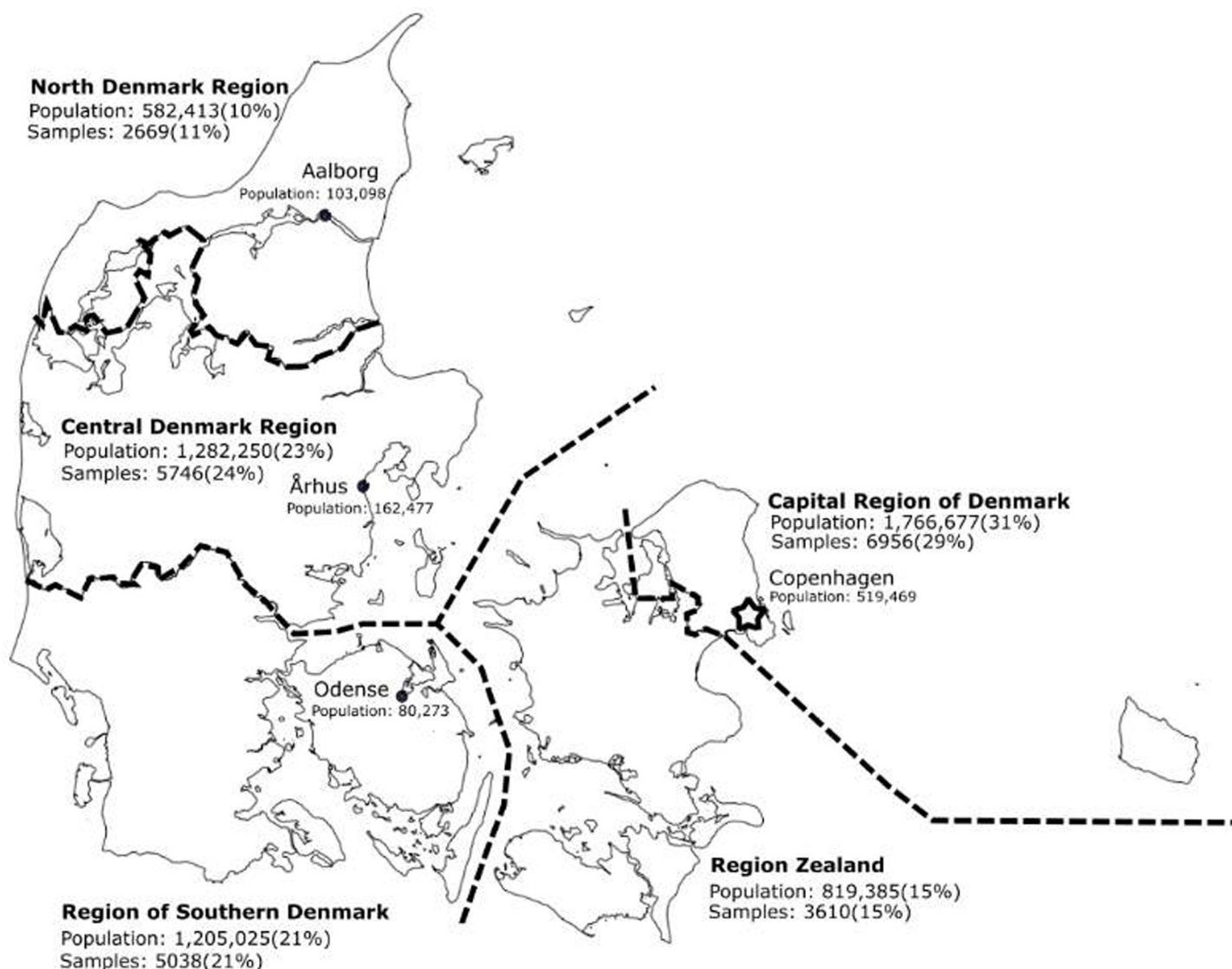


Fig 2. Map of Denmark with major administrative regions and metropolitan areas. The population of each region or area, as well as the number of samples, is given on the map. Based on birth-locations with ≥ 20 included individuals.

<https://doi.org/10.1371/journal.pone.0208829.g002>

Table 1. Distribution of mtDNA macro hgs L0-L6, M, N and R. (see [Materials and Methods](#)).

Macro-hg	n	%
L0	26	0.1
L1	13	0.1
L2	54	0.2
L3	71	0.3
L4	16	0.1
L5	2	0.0
L6	0	0.0
M	394	1.6
N	1,700	7.0
R	21,940	90.6
Total haplotyped	24,216	100

<https://doi.org/10.1371/journal.pone.0208829.t001>

Likewise, the median-joining graph of the R macro-hg (Fig 5), based on all the called SNPs, disclosed a phylogenetic relationship compatible with that of Fig 1. In addition, all the hgs exhibit a considerable complexity, suggesting that each hg is the result of multiple migratory events and thus, multiple founder events. The sub-hg distribution for each of the hgs that belong to the R macro-hg is shown in Table 3.

The L, N and M macro-hgs were infrequent (Table 1). M and N could be broken down into hgs as shown in Table 4. The complexity of these macro-hgs, and macro-hg L, as demonstrated by median-joining phylogenetic analysis, S1A–S1C Fig and PCAs S2A–S2C Fig, suggest that these macro-hgs are composed of specific haplogroups from multiple immigrations over time. The number of DNA variations between individual haplotypes in each haplogroup (S1A–S1C Fig) is so large that they cannot represent development from a single founder within the time frame defined by the length of time in which Denmark has been populated.

Geno-geographical affinity of mtDNA haplogroups

An admixture analysis of the persons of different hgs was performed with results as shown in Fig 6. The major hgs, H and its sub-hgs, have a 90–95% Danish ancestry and ~ 5% non-Danish European structure. However, there is a great variation between mtDNA hgs—most pronounced in hg U –, and most of the hgs have a notable—but varying—proportion of admixture from Europe, Middle East and Central South Asia (Fig 6). This finding is compatible with the very complex M-J diagram of the R-hgs (Fig 5). The genomic ancestry of the macro-hgs M and N differ even more (Fig 6), with the M being of South East and East Asian affinity, and N of Danish and European affinity. Macro-hg L exhibited—surprisingly—a predominance of Middle Eastern and European genomic affinity (Fig 6), where an African ancestry should be expected[39, 41].

Spatial distribution of mtDNA haplogroups

Denmark is divided into five geographical and administrative regions as shown in Fig 2. The samples for this study were obtained from all regions in Denmark and linked to the postal code of the birthplace. The frequency of the macro-hgs and the most frequent hgs is shown for each administrative region in Table 5. The most marked differences were the relatively low frequency of hg H, 43.4%, in the Capital Region, compared to the North Denmark Region with 48.8% hg H ($p < 0.0001$), and the higher frequencies of hgs L, M and R (in total 7.5%) in the Capital Region compared to the North Denmark Region (in total 3.9%) ($p < 0.0001$). These differences were significant even after Bonferroni correction for the simultaneous comparison of 14 groups

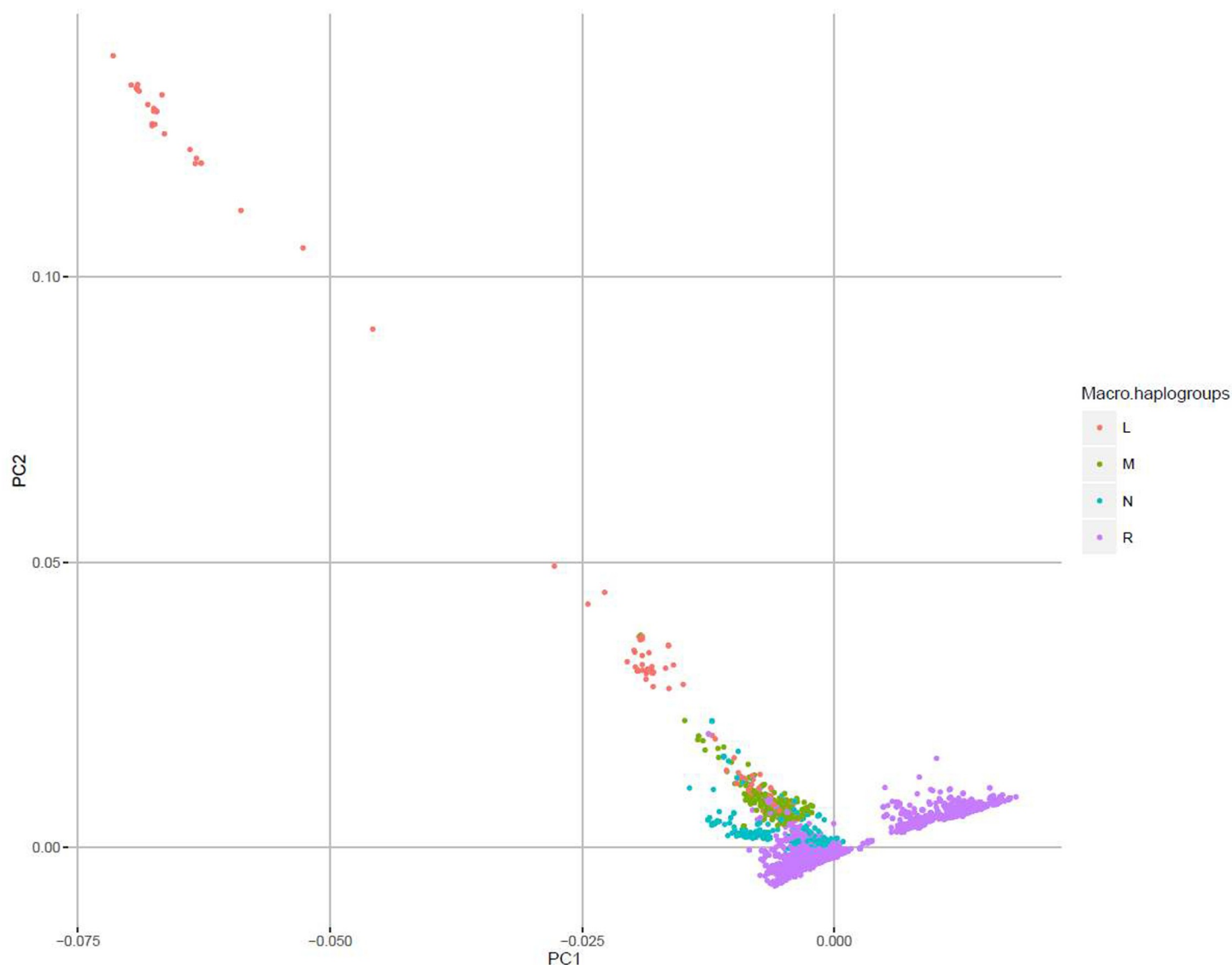


Fig 3. PCA of the macro-haplogroups L, M, N and R. PC1: First principal component. PC2: Second principal component.

<https://doi.org/10.1371/journal.pone.0208829.g003>

Table 2. Distribution of mtDNA haplogroups constituting the macro-hg R.

Hg	n	%
H	10,912	50.0
U	3,237	14.8
T	2,323	10.6
J	2,195	10.1
K	1,754	8.0
V	785	3.6
R	570	2.6
F	50	0.2
B	10	0.0
Total haplotyped	21,836	100

<https://doi.org/10.1371/journal.pone.0208829.t002>

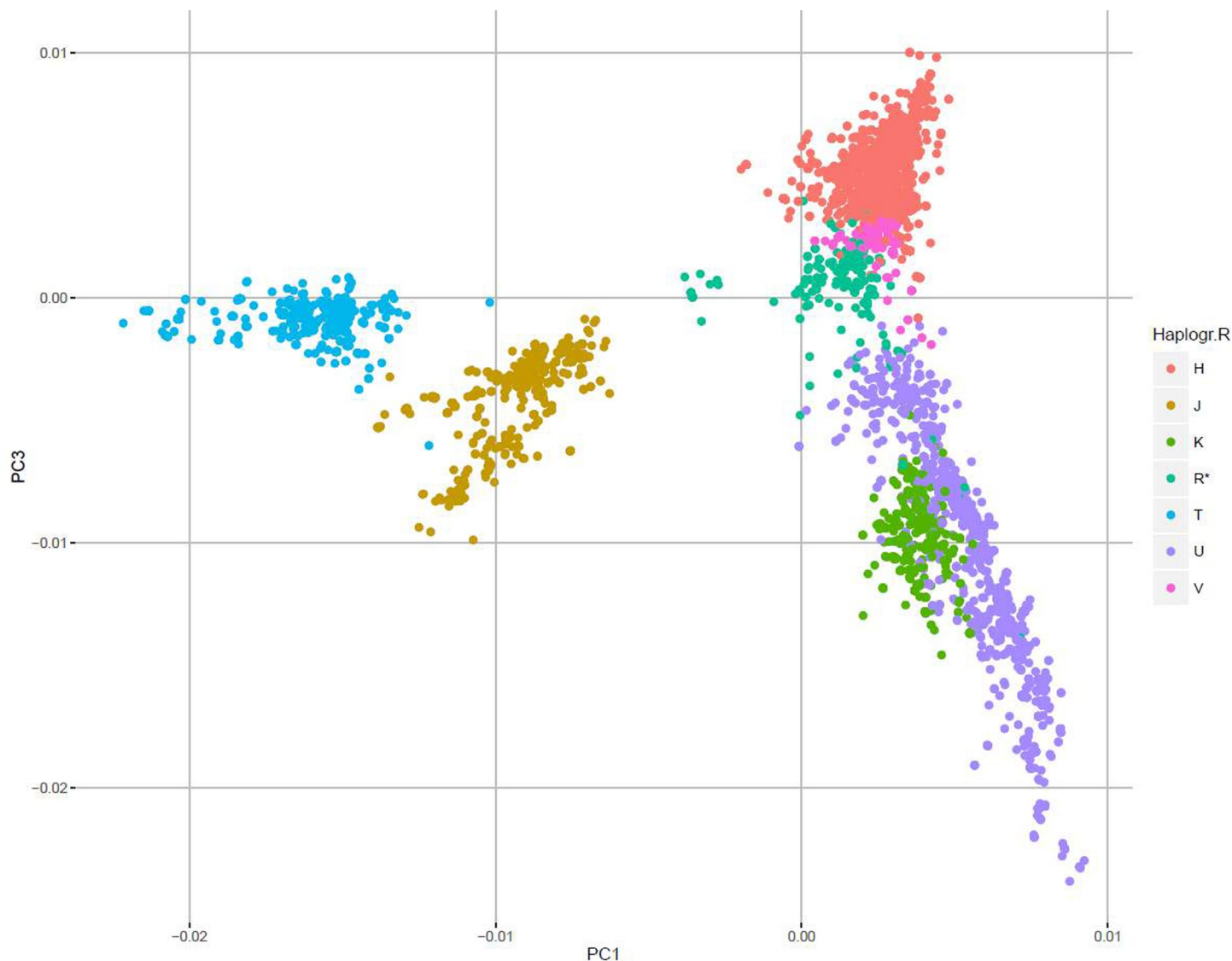


Fig 4. PCA of the macro-haplogroup R (the major European haplogroups). PC1: First principal component. PC3: Third principal component.

<https://doi.org/10.1371/journal.pone.0208829.g004>

(Table 5). In the Capital Region, the L hgs have a frequency of 1.3%, compared with 0.4–0.6% in the other regions and the M hgs have a frequency of 2.6% in the Capital region compared to 1.0–1.6% in the other regions. As the L and M hgs are rare in the European population and very frequent in African and Asian populations, the noted difference probably reflects a higher proportion and preferential localization of non-ethnic Danes in the Capital region. A similar spatial difference was apparent when the hg distributions of persons from the Danish metropolitan areas, comprising the major cities, Copenhagen, Aarhus, Odense and Aalborg, were compared with the distributions in the remaining rural areas (S4 Table). In the metropolitan areas, the combined frequency of L and M hgs was 4.4% as compared to 1.9% in the rural areas.

Temporal distribution of mtDNA haplogroups

The frequency of the major hgs H, J, T, K, V did not change significantly from year to year in the period from 1981 to 2005 (Results not shown).

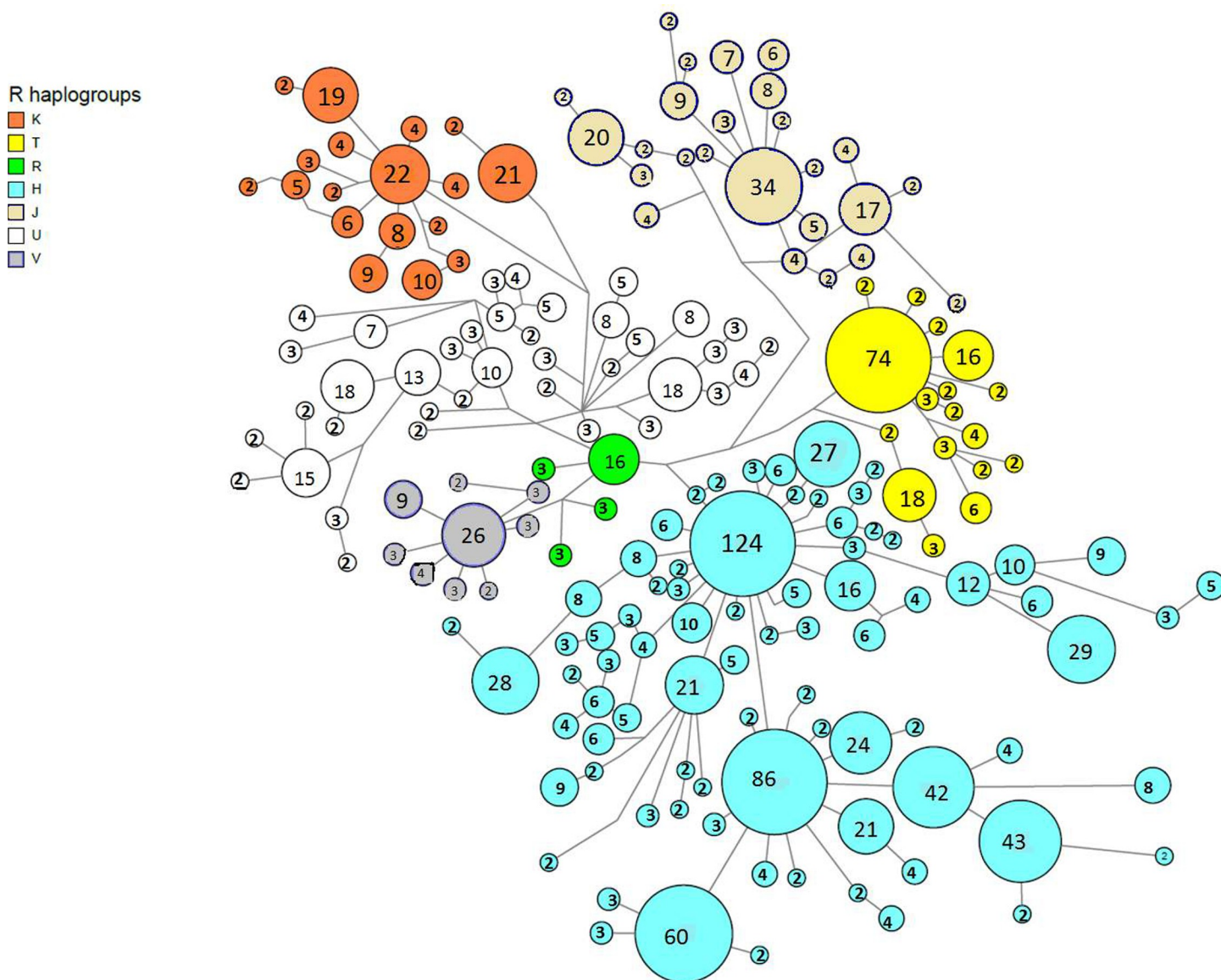


Fig 5. Median-joining (M-J) network of the macro-haplogroup R mtDNA sequences (2000 chosen at random).

<https://doi.org/10.1371/journal.pone.0208829.g005>

The frequencies of M and L hgs (Fig 7) increased over the period. The L hgs increased from a constant level of ~ 0.4% from 1981–1995 to ~ 1.5% in 2005. The M hgs rose from a constant level of ~ 1% from 1981–1991 to ~ 3% in 2005. The proportion of L and M hgs increased significantly (even after Bonferroni correction for the simultaneous analysis of 14 groups) from the period 1981–1986 to the period 2000–2005, Table 6. The considerable diversity of the M macro-hg (S1B Fig) and the extreme diversity of the L macro-hg (S1C Fig) as disclosed by the M-J-network, suggest that the hgs are the result of immigration from different source populations. The occurrence of new clusters on the PCA when the macro-hgs M and L from the period 1981–1986 were compared with the same macro-hgs in 2000–2005 (S3A and S3B Fig) makes it likely that the increase in the proportion of both macro-hgs represent immigration.

The hg U also increased in proportion each year from 1981 to 2005 albeit not significantly, following Bonferroni correction (Table 6). However, when analyzing only the R-macro-hg, the proportion of hg-U increased significantly ($p < 0.002$) (Fig 7). In the case of the U-hg, the PCA

Table 3. Distribution of mtDNA sub-haplogroups of H, V, J, T, K, and U, the most frequent European haplogroups. HV could not be defined and is included in the H hg.

Hg	Sub-hgs	n	%
H	H	3,924	36.4
	H1-H30b-H79a	4,054	37.6
	H2	1,058	9.8
	H3	699	6.5
	H4	220	2.0
	H5-36	828	7.7
	Number sub-haplotyped	10,783	100
U	U	120	3.8
	U1	66	2.1
	U2-U3	596	18.8
	U4-9	603	19.0
	U5a	1,113	35.1
	U5b	471	14.8
	U6	26	0.8
	U7	87	2.7
	U8	90	2.8
	Number sub-haplotyped	3,172	100
K	K	1	0.1
	K1	1,385	81.8
	K2	307	18.1
	K3	1	0.1
	Number sub-haplotyped	1,694	100
J	J	2	0.1
	J1	1,686	80.5
	J2	407	19.4
	Number sub-haplotyped	2,095	100
T	T*	9	0.4
	T1	396	17.1
	T2	1,908	82.5
	Number sub-haplotyped	2,313	100

<https://doi.org/10.1371/journal.pone.0208829.t003>

did not reveal the appearance of novel clusters when comparing 1981–1986 with 2000–2005 (S3C Fig). Thus, there was no evidence for the introduction of novel U-hgs over the period. However, a more detailed analysis revealed that the increase in U-hgs was largely due to an increase in the infrequent sub-hgs U1, U6, U7, and U8 (S3 Table) as all these increased by more than 40%. An admixture analysis showed that whereas the Danish autosomal genomic ancestry of the hgs U*, U2-U3, U4-9, U5a&b is in the range of 80.6% - 92.5% (Fig 6) comparable to that of the other major European hgs, the Danish ancestry drops to 36.4% - 42.3% for U1, U6 and U7. As illustrated in Fig 6, the U1 and U7 have strong Middle Eastern (29.4% and 13.6%) and Central South Asian (14.7% and 39.8%) genomic ancestry, and the U6 exhibits a very strong Middle Eastern genomic ancestry (53.8%). This suggests that the increase in frequency of hg U is largely due to expansion of hgs brought into Denmark as a result of recent immigration.

Discussion

This study shows that the distribution of mtDNA hgs in Denmark is highly dynamic and complex. It comprises 1.6% of the Danish population over a 25-year period, and is by far the largest

Table 4. Distribution of mtDNA haplogroups contained within the N and M macro hgs.

Macro-hg	(Sub) hgs	n	%
N	A	119	7.1
	I	655	39.2
	N	55	3.3
	N1-5	111	6.6
	N2	11	0.7
	W	319	19.1
	X	394	23.6
	Y	6	0.4
	Number haplotyped	1,670	100
M	C	40	10.7
	D4	53	14.2
	D5	4	1.1
	D6	2	0.5
	E	8	2.1
	G	10	2.7
	M	81	21.7
	M1	30	8.0
	M2	11	2.9
	M3	24	6.4
	M30	14	3.7
	M4	3	0.8
	M5	26	7.0
	M6	9	2.4
	M7	37	9.9
	M8a	7	1.9
	M9	4	1.1
	Z	11	2.9
	Number haplotyped	374	100

<https://doi.org/10.1371/journal.pone.0208829.t004>

performed of the distribution of mtDNA hgs in any country. The method of collecting stored DBS from the PKU biobank, where the coverage is ~ 99%[42], enabled us to survey a true population based sample of mtDNA from persons, where the time and place of birth was known from national electronic registries. This is in contrast to most other population genetic studies of adults sampled in a specific bias-prone context, e.g. hospitalized patients or geographically biased samplings.

The distribution of macro-hgs (Table 1) is typical of Northern Europe [43, 44]; 90.6% of individuals belonged to the R macro-hg. Furthermore, the distribution of hgs within the R macro-hg (Table 2) is very similar to that previously described among 9000+ persons from the greater Copenhagen area[45] and a much smaller (~200 cases) Danish forensic control sample [46], as well as that found in 2000 Danish exomes from a mixed control and diabetes population[47]. Additionally, The sub-hg distribution within macro-hg R, as presented in Li et al. [47], is similar to that seen in the present study (Table 3) e.g. H1 and H3 constitute 37.6% and 6.5% of the H hgs in our study and 39.6% and 5.2% in the Li et al. study[47].

Haplotyping of mtDNA was based on array data with only 418 mtDNA SNPs and as a consequence not all sub-hgs could be called. Ideally, sub-haplotyping could be refined—without including sub-hg defining SNPs from Phylotree—by the application of clustering approaches. However, this was not attempted, as it would lead to results that could not be compared with

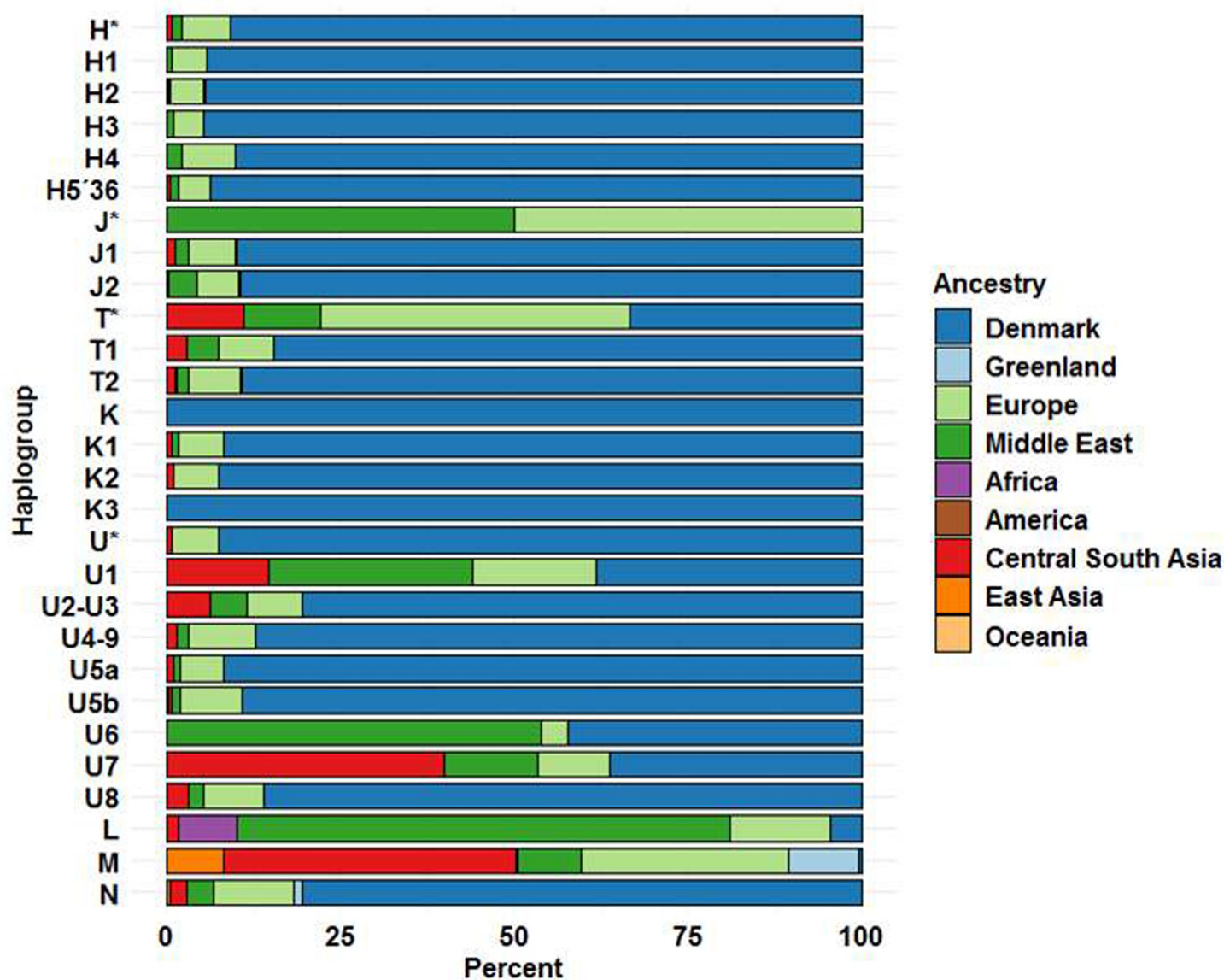


Fig 6. Genomic ancestry of samples as a function of mtDNA hg. (See [Materials and Methods](#) for details).

<https://doi.org/10.1371/journal.pone.0208829.g006>

other studies. The stringent adherence to specific SNPs meant that a number of persons, normally in the order of 1–2%—albeit somewhat higher for specific hgs (see [S2 Table](#)), were not assigned to a specific hg or sub-hg. An advantage of using a limited set of markers is that confounding due to private variants is avoided.

The complexity of each of the major hgs, as seen from the M-J-networks (Figs 5 and [S1A–S1C](#)) with multiple nodes and a span of many mutations between different leaves suggests that the hgs are not representative of single early founder events, as one mtDNA mutation is expected to occur per 3,000 years [48]. A notable exception is hg A, with a total prevalence of ~ 0.5% ([Table 4](#)) where the largest node ([S1A Fig](#)) is associated with “daughter” nodes only 1–2 mtDNA mutations from the major node. It was only possible to identify the sub-hgs A1 and A5, and they only constituted a small fraction of the total hg A. A likely source of the hg A is the Inuit population from Greenland, which is now a self-ruling part of

Table 5. Distribution of the most frequent haplogroups in the five administrative regions of Denmark. Only persons born at locations with ≥ 20 births are included.

Haplogroup	Capital Region	Region Zealand	Central Denmark Region	Region of South Denmark	North Denmark Region
	% (n)	% (n)	% (n)	% (n)	% (n)
H	43.4 (2,962)	46.9 (1,664)	45.7 (2,584)	45.6 (2,259)	48.8 (1,284)
I	2.6 (175)	2.5 (89)	3.2 (180)	3.0 (148)	2.3 (60)
J	8.6 (586)	8.8 (313)	9.4 (534)	9.4 (464)	8.2 (215)
K	6.9 (468)	6.3 (224)	8.1 (456)	7.3 (364)	8.7 (229)
L	1.3 (88)	0.5 (17)	0.5 (31)	0.4 (20)	0.6 (17)
M	2.6 (176)	1.3 (45)	1.0 (56)	1.6 (79)	1.0 (27)
N	1.3 (92)	1.2 (43)	1.2 (68)	1.2 (61)	1.3 (33)
R	3.6 (244)	2.1 (74)	2.5 (144)	2.4 (119)	2.3 (61)
T	9.2 (627)	10.0 (354)	9.3 (525)	8.8 (434)	9.0 (238)
U	13.8 (939)	13.8 (491)	13.1 (744)	13.4 (666)	13.1 (346)
V	3.4 (229)	3.3 (116)	3.4 (195)	3.5 (175)	2.8 (73)
W	1.4 (96)	1.6 (57)	1.3 (74)	1.4 (67)	0.9 (24)
X	2.0 (134)	1.7 (60)	1.2 (67)	2.0 (101)	0.9 (24)
NA	0 (2)	0.1 (3)	0 (2)	0 (1)	0 (1)
Total:	100 (6,818)	100 (3,550)	100 (5,660)	100 (4,957)	100 (2,632)

<https://doi.org/10.1371/journal.pone.0208829.t005>

the Kingdom of Denmark. Several studies have established that the major hg in Greenlandic Inuits is hg A2 followed by hg D3 [49–52].

The lack of recombination, high mutation frequency, and fixation of mtDNA hgs has enabled the use of mtDNA in population genetics to study population ancestry, migrations, gene flow and genetic structure [53, 54]. Thus, populations on different continents, e.g. Native Americans [55], Africans [56] and Europeans [57] were ascribed specific matrilineal mtDNA hg distributions [39, 44]. However, recent studies using autosomal SNP markers have disclosed a considerable ancestral complexity underlying an mtDNA classification in specific admixed populations, and the prediction of a specific mtDNA hgs is not possible from a specific continental ancestry based on nuclear genetic markers [58]. A problem with genetic analysis of admixed populations is the lack of temporal resolution. It is often not possible to date a specific population split because differentiation between new and old migrations is impossible. Recent advances in sequencing ancient genomes [59] have made it possible to combine genetic information from ancient humans with archaeological information on the age of skeletal remains [60], thus constructing hg distribution maps with a temporal dimension for e.g. Ice age and Bronze age Europe [61, 62]. Such maps may help explain demic and cultural exchange [52, 63–65].

There is no solid evidence of the presence of humans in Denmark [66] [67] prior to the Last Glacial Maximum (LGM) (26.5–19 kYBP) [68]. At that time Denmark was covered in ice, except for the south-western part of Jutland [69], and following the retraction of the ice sheet, peopling became possible from the south [69–72]. The first inhabitants documented were late Paleolithic hunters entering from southern Europe [73]. These hunters are discernible from Bølling time [69] around 12,800 BC. The archaeological remains suggest their transient presence in seasonal hunting periods until the Mesolithic around 9,700 BC [74]. The Hamburgian, Federmesser, Bromme and Ahrensburgian material cultures [69, 75], well known from findings in Germany [69], are represented during this period. The earliest anatomically normal humans (ANM), present from around 45–41.5 kYBP [76–78], where ancient DNA studies have revealed the presence of mtDNA hgs M and pre-U2 [79], have little similarity to present-day Europeans. However, the Europeans from around 37 kYBP to 14 kYBP have left their

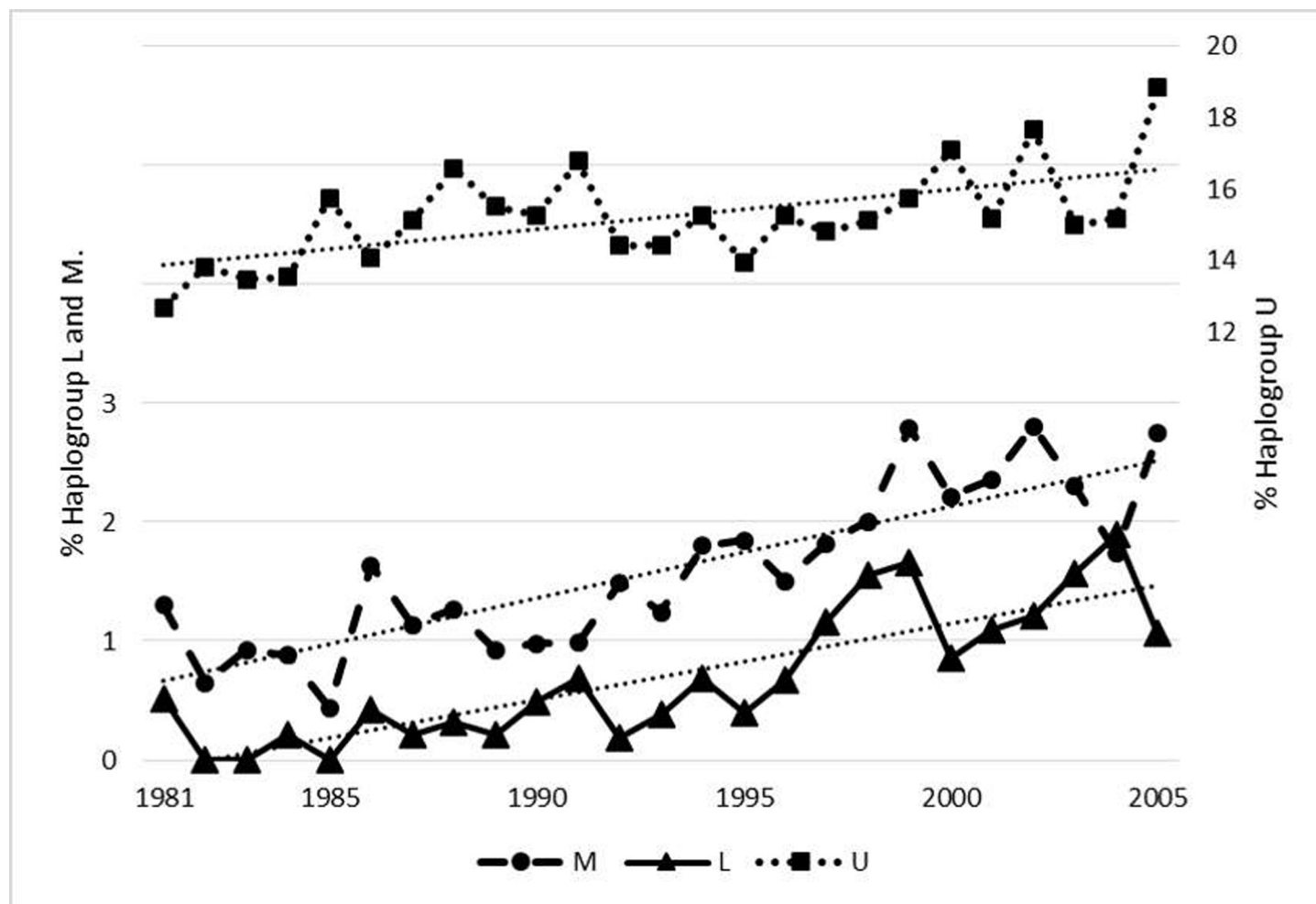


Fig 7. The proportion of haplogroup L, M and U as a function of birth year.

<https://doi.org/10.1371/journal.pone.0208829.g007>

mark on present-day Europeans [62]. From 14 kYBP the European population has a strong near-eastern component [62]. Around 7 kYBP the Neolithic transformation gradually started as the result of a demic dissemination of Neolithic Aegeans [80]. Thus, a minimum of three ancestral populations, i.e. a western European hunter-gatherer, an ancient north Eurasian, and an early European farmer population are needed to explain present-day European autosomal genome compositions [81]. In Denmark, where Paleolithic ancient genetic data have not been published, ancient mtDNA haplotyping of three Neolithic corpses from 4.2–4 kYBP revealed two U4 and one U5a mtDNA hgs [82] and later samples showed a mixture of hgs also found in northern Germany [79, 82]. The presence of mtDNA clades deriving from Paleolithic and Neolithic Europeans in the extant Danish population is thus explained.

In the Bronze age, an influx of people from the Russian steppe and North Caucasus, bringing the Indo-European language and culture [83], resulted in the last major prehistoric demic change [61] in Europe. In historic times the migrations have been many, particularly in the half-millennium following the fall of the Western Roman Empire [84]. In Denmark, apart from continuous demic exchange with Southern Scandinavia and present-day Germany [69, 72], early historic time was mostly characterized by emigration, i.e. the Heruli, Cimbriae and Teutones, Burgundians and Vikings [71, 72]. The first census held in 1769 AD reported the

Table 6. Distributions of the most frequent haplogroups in the period 1981–1986 and 2000–2005.

Haplogroup	1981–1986	2000–2005	P-Value
	% (n)	% (n)	
H	47.3 (2,162)	44.3 (2,454)	2.98E-03
I	2.8 (130)	2.7 (150)	0.71
J	8.9 (405)	9.2 (511)	0.53
K	7.9 (359)	7.0 (387)	0.10
L	0.2 (8)	1.2 (67)	1.14E-10
M	1.0 (44)	2.4 (131)	3.73E-08
N	1.1 (50)	1.3 (73)	0.32
R	2.2 (102)	2.9 (163)	0.03
T	9.8 (447)	8.5 (468)	0.02
U	12.7 (579)	14.3 (790)	0.02
V	3.6 (163)	3.1 (169)	0.16
W	1.2 (56)	1.4 (80)	0.39
X	1.4 (64)	1.6 (89)	0.41
NA	0.0 (1)	0.1 (3)	0.63
Total	100 (4,570)	100 (5,535)	-

<https://doi.org/10.1371/journal.pone.0208829.t006>

population of what is present day Denmark to be ca 800,000 persons [85]. This number has increased to 5.6 mio in 2014, despite emigration of 287,000 persons between 1867 and 1914 (~10% of the population) largely to Northern America, and the same number from 1914 to 1968 [86]. After the Second World War, Denmark has seen considerable immigration from other European countries, but also from Asian, Middle Eastern and African countries (S4 Fig). This part of Danish history can explain the occurrence of a diversity of mtDNA clades deriving from a large number of geographic regions. The extant distribution of mtDNA hgs reported here (Tables 1–4) reflect the complicated and heterogeneous origins of the people currently inhabiting Denmark.

The distribution of H sub-hgs (Table 3) and the M-J graph of the H-hg (Fig 5) with multiple major nodes and 5–10 variants between leaves, suggest that the H-hg lineages are the result of repeated immigrations. Most likely from northern Europe, where a study of 39 prehistoric hg H samples has shown that the distribution of hg H differed between early and middle-to-late Neolithic groupings. Prior to Neolithicum H hgs have not been found in skeletal remains; an ensemble of Swedish Mesolithic hunter-gatherers all had U-hgs [63]. Whereas H, H5 and H1 were found throughout Neolithicum, H5b, H10, H16, H23, H26, H46, H88, and H89 were seen in early Neolithic samples, and H2, H3, H4, H5a, H6, H7, H11, H13, H82, and H90 in middle-to-late Neolithic samples [87]. The extant Danish H-hg distribution is compatible with contributions from throughout Neolithicum, but estimating when the hgs appeared in Denmark remains unfeasible, as the distribution is similar to that of northern Germany. However, not all carriers of an H-hg have a Danish autosomal genomic ancestry (Table 5) compatible with a considerable recent immigration from countries with a European mtDNA distribution (S4 Fig).

The most frequent Danish U-sub-hg is U5 (Table 3) which is an old European mtDNA hg, with two major subclades (U5a and U5b) with coalescence time estimates of 16–20 kYBP and 20–24 kYBP, respectively [88]. The U5-hg is the most frequent U-sub-hg after LGM [79] and the carriers have a Danish autosomal genomic ancestry around 90%, and a European ancestry of 6.1–8.8% (Fig 6). However, several of the less frequent U-sub-hgs have a considerable, ~40–60% non-Danish and non-European genomic ancestry based on autosomal markers, with U7

related to South East Asia and U6 to the Middle East, respectively. In addition, J & T hgs (Fig 6) exhibit a relatively strong non-Danish autosomal genomic ancestry. The M-J graphs (Fig 5) also disclose a considerable variation, much larger than could be attained merely within the timeframe where Denmark has been populated. This may be explained by an extensive immigration, which for U6 and U7 occurred recently. This is also compatible with the rising frequency of these hgs during the 25-year study period (Fig 7 and S3 Table).

The N-macro-hg exhibits a high (>90%) combined Danish and European genomic ancestry (Fig 6) suggesting that the major part has been in Denmark for a long time. The major N-hgs are the I-hg, which has been found in meso- and neolithic Scandinavians [82] and hg-X [89] and hg-W (Table 4) and they exhibit (S1 and S2A Figs) a very extensive heterogeneity. All three hgs are old and have a broad, low frequency, distribution in western Eurasia [43], resulting from migratory events from the Near East and Central Asia. These events, viz. the significance for the presence of the hgs in Denmark, cannot be temporally resolved.

Macro-hg-M, while infrequent (1.6%) (Table 1) has also increased in frequency recently (Fig 7) and exhibits extensive heterogeneity (S1 Fig) in the M-J graph and the PCA (S2B Fig) suggesting many migratory events of diverse origins. The largest group is M, that could not be decomposed further, but the second most frequent, D4 (14.2% of M-hgs), is most likely of central Asian origin [90], and the M2, M3, M4 and M6 hgs, constituting 12.5% of M-hgs, are of Indian or Pakistani origin [91]. These findings are compatible with the high genomic ancestry (42.2%) (Fig 6) towards Central South Asia, and a recent entry to Denmark. There is a propensity for location in Metropolitan areas (S4 Table) and the Capital region (Table 5) also compatible with recent immigration.

The low, but increasing, frequency of L-hgs (Tables 1 and 6 and Fig 7) predominantly L2 and L3 (Table 1), is also the result of multiple immigration events, as the PCA (S2C Fig) as well as the M-J plot reveals an extensive heterogeneity (S1C Fig). L2 and L3 hgs are frequent in south, west or east Africa, whereas the contribution from central Africa is small [92, 93]. However, the admixture analysis (Fig 6) suggests a much stronger genomic ancestry with the Middle East (70.8%) than to Africa (8.4%). However, this finding could be an artefact of the ADMIXTURE analysis, as the highly variable African genomic ancestry [41] is represented by only 127 genomes (~ 6% of the total reference set—see Materials and Methods), as compared to 178 genomes from the Middle East, making the definition of African ancestry imprecise. Alternatively, it could be caused by the extensive demic exchange that have occurred through time between the Near East and Northern Africa [85, 94]. The number of children born in the period 1980 to 2005 with one or two African parents (See S4 Fig) is compatible with the origin of the L hg persons being African. The L hgs are predominantly located in the Capital Region (Table 5) and in Metropolitan areas (S4 Table) which is also compatible with recent immigration.

Immigration to Denmark increased from 1980, where 135,000 immigrants were registered, to 2005, where this number had risen to 345,000 [85]. In the same period the number of descendants of immigrants, i.e. persons that might turn up in this study, rose from 18,000 to 109,000. Roughly 50% of the immigrants were from western countries [85] (S4 Fig). This should give approximately 45,000 descendants of non-western immigrants over the time of study. This is within the order of magnitude to be expected from the frequency distribution of mtDNA hgs in Denmark. There is thus a reasonable concordance between the suspected number of immigrants, from the temporal and structural study of mtDNA hgs, and the registered births of persons with non-Danish parents.

Whereas the temporal change in mtDNA distribution can be explained by immigration, it is more difficult to explain the spatial clines (Tables 5 and S4). The frequency of hg-H is higher in Northern Jutland than in other regions, particularly the Capital Region (Table 5). The

difference cannot be explained by the comparatively much smaller differences in frequencies of hgs L and M. As the mobility of Danes was fairly restricted until around 1900 [71], it may represent differences caused by centuries of relative isolation of a population north of Limfjorden. In Slovenia, historically confirmed geographically-based sub-stratification of the population [95] has led to extreme differences in mtDNA distributions.

A recent fine-scale gDNA population structure study from the UK [37] revealed considerable geographical heterogeneity and enabled the identification of specific sources of admixture from continental Europe. A gDNA study of admixture in the Danish population [96] showed considerably more homogeneity, but medieval admixture from Slavic tribes in North Germany, as well as a North-South gradient were discernible. Both of these studies limited the participants to persons with local grandparents, whereas our present study was not directed towards previous generations, but rather present-time and prospective.

The association of mtDNA SNPs and hgs with both diseases and functional characteristics of mitochondria, has led to a pathogenic paradigm [97] where variation in mitochondrial function is considered to be of paramount importance for development of disease. Specific hgs have also been associated with longevity [98] and likelihood of being engaged in endurance athletic activities [99]. The clinical presentation of diseases caused by specific mtDNA variants depends, in some cases, on the hg background [100]. However, several of these studies are underpowered [101], poorly stratified with respect to sex, age, geographical background [102] or population admixture [103], or have used small areas of recruitment risking “occult” founder effects [104]. Consequently, there are many contradictory results reported in the literature. To circumvent some of these problems a recent large study on mtDNA SNPs identified a number of SNPs that were associated with several degenerative diseases [105], however, the study pooled sequence information from a large geographical area, without correcting for potential population sub-structure.

In addition to the gDNA/mtDNA interaction (Fig 6) the demonstrated spatio-temporal dynamics of the mtDNA hg distribution (Table 5 and Fig 7) should be taken into account when designing studies of mtDNA associations with disease, physiological characteristics and pharmacological effects. A prerequisite for genetic association studies is that the individuals are sampled from a homogenous population, or that cryptic population structure is corrected for, if not, false positive associations may occur [106–108]. It is even more complicated, as a functional interaction between different mtDNA hgs and nuclear genomes, of importance for longevity and insulin resistance, has been described in mice [109]. Thus, the combined variation in mtDNA and gDNA should be accounted for in association studies. This may be done by including the principal components of both in the association analysis.

The need for well-designed and robust association studies in mitogenomics is underscored by the difficulty of performing functional analyses of mtDNA variants. Thus, OXPHOS enzyme activity measurements in cells and tissues [110], bioinformatic analysis of the consequences of altered mitochondrial protein function [111], as well as cybrid studies [112] are all characterized by not taking the mitochondrial-nuclear interaction into account. This is a major problem as changes in mitochondrial function can influence a plethora of nuclear functions, and vice versa [113].

Thus, when studying bi-genomic, i.e. both nucleo- and mito-genomic, disease associations, our results suggest that it is necessary to compensate for gDNA and mtDNA genetic stratification, as well as the interaction between the two sources of variation and spatio-temporal clines, in order to establish a functional significance of a specific mtDNA hg or SNP. To our knowledge, this has never been done, suggesting that previous reports on disease associations and mtDNA haplogroups should be considered preliminary.

Materials and methods

Ethics statement

This is a register-based cohort study solely using data from national health registries. The study was approved by the Scientific Ethics Committees of the Central Denmark Region (www.komite.rm.dk) (J.nr.: 1-10-72-287-12) and executed according to guidelines from the Danish Data Protection Agency (www.datatilsynet.dk) (J.nr.: 2012-41-0110). Passive consent was obtained, in accordance with Danish Law nr. 593 of June 14, 2011, para 10, on the scientific ethics administration of projects within health research. Permission to use the DBS samples stored in the Danish Neonatal Screening Biobank (DNSB) was granted by the steering committee of DNSB (SEP 2012/BNP).

Persons

As part of the iPSYCH (www.iPSYCH.au.dk) recruitment protocol, 24,651 singletons (47.1% female), born between May 1 1981 and Dec 31 2005 were selected at random from the Danish Central Person Registry. The singletons had to have been alive one year after birth, and to have a mother registered in the Danish Central Person Registry. Furthermore, it should be possible to extract DNA from the DBS. DBS cards were obtained from the Danish Neonatal Screening Biobank at Statens Serum Institut [114] and DNA was extracted and analyzed as described below. At the time of analysis (2012), the mean age of females was 18.2 years (SD: 6.6 years) and for males 18.8 years (SD: 6.7 years). There was no bias in the geographical distribution of the birthplace of samples (Fig 2).

Genetic analysis

From each DBS card two 3.2-mm disks were excised from which DNA extracted using Extract-N-Amp Blood PCR Kit (Sigma-Aldrich, St Louis, MO, USA) (extraction volume: 200 μ L). The extracted DNA samples were whole genome amplified (WGA) in triplicate using the REPLig kit (Qiagen, Hilden, Germany), then pooled into a single aliquot. Finally, WGA DNA concentrations were estimated using the Quant-IT Picogreen dsDNA kit (Invitrogen, Carlsbad, CA, USA). The amplified samples were genotyped at the Broad Institute (MA, USA) using the Psychiatric Genetic Consortia developed PsychChip (Illumina, CA, USA) typing 588,454 variants. Following genotyping, samples with less than 97% call rate, as well as those where the estimated gender differed from the expected gender were removed from further analysis; altogether 435 (1.8%) samples were removed due to problems with calling mtDNA variants. We then isolated the 418 mitochondrial loci and reviewed the genotype calls, before exporting into the PED/MAP format using GenomeStudio (Illumina, CA, USA). Samples were loaded into GenomeStudio (version 2011.a), a custom cluster was created using Gentrain (version 2), following automatic clustering all positions with heterozygotes were manually curated. The custom cluster file was trained on genomic DNA and optimized for clustering of AA, AB, and BB. As mtDNA exhibits heteroplasmy each cluster was manually inspected and adjusted based on Norm R (intensity) and Norm Theta (allele frequency). Samples with low intensities (< 0.08) were considered unreliable calls and rejected. Norm Theta was used to adjust the clusters into two distinct clusters, AA and BB. The few samples that clustered in AB were regarded as missing calls. The data was exported relative to the forward strand using PLINK Input Report Plug-in (version 2.1.3). Eigenvectors were calculated using PLINK (v1.90b3.31). PCA plots were created using the package ggplot2 (version 1.0.1) in R (version 3.1.3). The call rate as a function of

the year the DBS were collected are given in [S5 Fig](#). There was no discernible effect of storage up to 25 years.

mtDNA SNPing

Haplotyping of mtDNA was performed manually using the defining SNPs reported in www.phylotree.org [19]. Hierarchical affiliation to macro-hg i.e. L0–L6, M, N, R, and subsequent to hgs—units more distal in the cladogram ([Fig 1](#)) was performed. In some cases it was possible to establish affiliation to even sub-hgs. The call efficiencies of SNPs used in defining haplo- and sub-haplogroup affiliation are summarized in [S1 Table](#). The proportion of samples that could not be unequivocally distributed into hgs or sub-hgs are given in [S2 Table](#), with the appropriate references to the relevant tables in the main manuscript.

Phylogenetic analyses

Phylogenetic analyses was performed by constructing median-joining networks with Network 4.6.1.3 (<http://www.fluxus-engineering.com>). FASTA converted sequences of mtDNA SNPs from each person were aligned, sequences were pre-processed with Star Contraction (Maximum star radius 5 for R, N, M and 1 for L), then Median Joining networks were constructed (using the network parameters: Epsilon 10, Frequency > 1, active) followed by post-processing with a maximum parsimony algorithm (MP) [115, 116]. Network Publisher were used to post-process the networks [42].

Genomic ancestry analysis

Ancestry estimation was done using ADMIXTURE 1.3.0 [117]. Briefly, a reference population consisting of Human Genome Diversity Project (HGDP) (<http://www.hagsc.org/hgdp/>) genotyping SNP data set, supplemented with representative samples of Danes (716 individuals) and Greenlanders (592 individuals) available at SSI from unrelated projects, was used. The final reference data set consisted of 103,268 SNPs and 2,248 individuals assigned to one of nine population groups: Africa, America, Central South Asia, Denmark, East Asia, non-Danish Europe, Greenland, Middle East and Oceania. K—number of clusters defined—was set to eight, based on principal component analysis clustering (data not shown).

Individuals characterized by different mtDNA hgs or sub-hgs were merged with the reference population data set and analyzed using ADMIXTURE. For prediction of the ancestry of individuals within the mtDNA hgs we created a random forest model [118] based on the reference data set, with the clusters Q1–8 as predictors and population groups as outcome. The prediction was thus supervised. Prediction was done in R version 3.2.2, using the caret package. The distribution of the eight basic clusters in samples of different geographical origin is shown in [S1 Fig](#). As expected, the African-characteristic cluster distribution plays a decreasing role when going from Africa over the Middle East to Central South Asia. Likewise, the Danish cluster distribution is very similar to that of Europe.

Statistics

The statistical significance of differences in mtDNA proportions was assessed using a permutation version of Fisher's exact test [119]. Calculations were performed using R [120]. To assess population stratification [121], principal component analysis (PCA) was performed. When evaluating the significance of multiple comparisons between two populations/groups, Bonferroni correction was used.

Supporting information

S1 Fig. M-J Networks of mtDNA haplogroups. M-J Network of haplogroups belonging to A) N-macrohaplogroup, B) M macrohaplogroup and C) L macrohaplogroup. (TIF)

S2 Fig. PCAs of mtDNA SNPs. PCA of mtDNA SNPs from persons belonging to A) N macrohaplogroup, B) M macrohaplogroup and C) L macrohaplogroup. The affiliation to haplogroup is shown with color coding. (TIF)

S3 Fig. PCAs of A) hg L, B) hg M and C) hg U in 1981–1986 and 2000–2005, PC1: First principal component. PC2: Second principal component. (TIF)

S4 Fig. Cumulative number of children born of non-Danish citizens as a function of geographical origin. (TIF)

S5 Fig. The call-rate of SNPs as a function of the year of deposition of the DBS. (TIF)

S1 Table. Haplogroup and sub-haplogroup defining SNPs and call frequency. (DOCX)

S2 Table. Number and proportion of samples that could not be haplotyped at different steps of the haplotyping and sub-haplotyping. In brackets are given the tables that are referred to in the manuscript. (DOCX)

S3 Table. Proportion of different Hg-U subhaplogroups in 1981–1986 and from 2000–2005. U*: All U hgs not belonging to any of the specified U sub-hgs. (DOCX)

S4 Table. Relative proportions of macro-hgs L, M and U in metropolitan and rural areas in 1981–1986 and 2000–2005. (DOCX)

Acknowledgments

The iPSYCH study was funded by The Lundbeck Foundation Initiative for Integrative Psychiatric Research. This research has been conducted using the Danish National Biobank resource, supported by the Novo Nordisk Foundation.

Author Contributions

Conceptualization: Jonas Bybjerg-Grauholm, Vanessa F. Gonçalves, Paula L. Hedley, Jørgen K. Kanter, James Kennedy, Michael Christiansen.

Data curation: Jonas Bybjerg-Grauholm, Christian M. Hagen, Marie Bækvad-Hansen, Christine S. Hansen, Paula L. Hedley, Ole Mors, Thomas D. Als, Preben B. Mortensen, David M. Hougaard.

Formal analysis: Thomas D. Als, Alfonso B. Demur, Michael Christiansen.

Funding acquisition: Ole Mors, Merete Nordentoft, Anders Børghlum, Preben B. Mortensen, Thomas M. Werge, David M. Hougaard, Michael Christiansen.

Investigation: Jonas Bybjerg-Grauholm, Christian M. Hagen, Marie Bækvad-Hansen, Jimmi Nielsen, Michael Theisen, James Kennedy, Thomas D. Als, Alfonso B. Demur, Merete Nordentoft, Thomas M. Werge, Michael Christiansen.

Methodology: Jonas Bybjerg-Grauholm, Christian M. Hagen, Vanessa F. Gonçalves, Marie Bækvad-Hansen, Christine S. Hansen, Jimmi Nielsen, Michael Christiansen.

Project administration: Jonas Bybjerg-Grauholm, Ole Mors, Anders Børghlum, Preben B. Mortensen, Thomas M. Werge, Michael Christiansen.

Resources: Michael Christiansen.

Software: Christian M. Hagen, Thomas D. Als.

Supervision: Paula L. Hedley, Merete Nordentoft, Anders Børghlum, Preben B. Mortensen, Thomas M. Werge, David M. Hougaard, Michael Christiansen.

Validation: Vanessa F. Gonçalves.

Visualization: Michael Christiansen.

Writing – original draft: Michael Christiansen.

Writing – review & editing: Jonas Bybjerg-Grauholm, Christian M. Hagen, Vanessa F. Gonçalves, Marie Bækvad-Hansen, Christine S. Hansen, Paula L. Hedley, Jørgen K. Kanter, Jimmi Nielsen, Michael Theisen, Ole Mors, James Kennedy, Thomas D. Als, Alfonso B. Demur, Merete Nordentoft, Anders Børghlum, Preben B. Mortensen, Thomas M. Werge, David M. Hougaard.

References

1. Papa S, Martino PL, Capitanio G, Gaballo A, De Rasmo D, Signorile A, et al. The oxidative phosphorylation system in mammalian mitochondria. *Adv Exp Med Biol.* 2012; 942:3–37. https://doi.org/10.1007/978-94-007-2869-1_1 PMID: 22399416.
2. Antico Arciuch VG, Elguero ME, Poderoso JJ, Carreras MC. Mitochondrial regulation of cell cycle and proliferation. *Antioxid Redox Signal.* 2012; 16(10):1150–80. <https://doi.org/10.1089/ars.2011.4085> PMID: 21967640; PubMed Central PMCID: PMC3315176.
3. Carafoli E. The fateful encounter of mitochondria with calcium: how did it happen? *Biochim Biophys Acta.* 2010; 1797(6–7):595–606. <https://doi.org/10.1016/j.bbabi.2010.03.024> PMID: 20385096.
4. Zhang F, Zhang L, Qi Y, Xu H. Mitochondrial cAMP signaling. *Cell Mol Life Sci.* 2016; 73(24):4577–90. <https://doi.org/10.1007/s00018-016-2282-2> PMID: 27233501; PubMed Central PMCID: PMC45097110.
5. Chinnery PF, Hudson G. Mitochondrial genetics. *Br Med Bull.* 2013; 106:135–59. <https://doi.org/10.1093/bmb/ldt017> PMID: 23704099; PubMed Central PMCID: PMC3675899.
6. Andrews RM, Kubacka I, Chinnery PF, Lightowlers RN, Turnbull DM, Howell N. Reanalysis and revision of the Cambridge reference sequence for human mitochondrial DNA. *Nat Genet.* 1999; 23(2):147. <https://doi.org/10.1038/13779> PMID: 10508508.
7. Calvo SE, Mootha VK. The mitochondrial proteome and human disease. *Annu Rev Genomics Hum Genet.* 2010; 11:25–44. <https://doi.org/10.1146/annurev-genom-082509-141720> PMID: 20690818; PubMed Central PMCID: PMC3397899.
8. Pagliarini DJ, Calvo SE, Chang B, Sheth SA, Vafai SB, Ong SE, et al. A mitochondrial protein compendium elucidates complex I disease biology. *Cell.* 2008; 134(1):112–23. <https://doi.org/10.1016/j.cell.2008.06.016> PMID: 18614015; PubMed Central PMCID: PMC2778844.
9. Pyle A, Hudson G, Wilson IJ, Coxhead J, Smertenko T, Herbert M, et al. Extreme-Depth Re-sequencing of Mitochondrial DNA Finds No Evidence of Paternal Transmission in Humans. *PLoS Genet.* 2015; 11(5):e1005040. <https://doi.org/10.1371/journal.pgen.1005040> PMID: 25973765; PubMed Central PMCID: PMC4431825.

10. Brown WM, George M Jr., Wilson AC. Rapid evolution of animal mitochondrial DNA. *Proc Natl Acad Sci U S A*. 1979; 76(4):1967–71. PMID: [109836](#); PubMed Central PMCID: PMC383514.
11. Wallace DC, Singh G, Lott MT, Hodge JA, Schurr TG, Lezza AM, et al. Mitochondrial DNA mutation associated with Leber's hereditary optic neuropathy. *Science*. 1988; 242(4884):1427–30. PMID: [3201231](#).
12. Holt IJ, Harding AE, Morgan-Hughes JA. Deletions of muscle mitochondrial DNA in patients with mitochondrial myopathies. *Nature*. 1988; 331(6158):717–9. <https://doi.org/10.1038/331717a0> PMID: [2830540](#).
13. Bourgeron T, Rustin P, Chretien D, Birch-Machin M, Bourgeois M, Viegas-Pequignot E, et al. Mutation of a nuclear succinate dehydrogenase gene results in mitochondrial respiratory chain deficiency. *Nat Genet*. 1995; 11(2):144–9. <https://doi.org/10.1038/ng1095-144> PMID: [7550341](#).
14. Vafai SB, Mootha VK. Mitochondrial disorders as windows into an ancient organelle. *Nature*. 2012; 491(7424):374–83. <https://doi.org/10.1038/nature11707> PMID: [23151580](#).
15. Koopman WJ, Willems PH, Smeitink JA. Monogenic mitochondrial disorders. *N Engl J Med*. 2012; 366(12):1132–41. <https://doi.org/10.1056/NEJMr1012478> PMID: [22435372](#).
16. Wallace DC, Brown MD, Lott MT. Mitochondrial DNA variation in human evolution and disease. *Gene*. 1999; 238(1):211–30. PMID: [10570998](#).
17. Ruiz-Pesini E, Mishmar D, Brandon M, Procaccio V, Wallace DC. Effects of purifying and adaptive selection on regional variation in human mtDNA. *Science*. 2004; 303(5655):223–6. <https://doi.org/10.1126/science.1088434> PMID: [14716012](#).
18. Bowton E, Field JR, Wang S, Schildcrout JS, Van Driest SL, Delaney JT, et al. Biobanks and electronic medical records: enabling cost-effective research. *Sci Transl Med*. 2014; 6(234):234cm3. <https://doi.org/10.1126/scitranslmed.3008604> PMID: [24786321](#); PubMed Central PMCID: PMC4226414.
19. van Oven M, Kayser M. Updated comprehensive phylogenetic tree of global human mitochondrial DNA variation. *Hum Mutat*. 2009; 30(2):E386–94. <https://doi.org/10.1002/humu.20921> PMID: [18853457](#).
20. Chinnery PF, Taylor GA, Howell N, Andrews RM, Morris CM, Taylor RW, et al. Mitochondrial DNA haplogroups and susceptibility to AD and dementia with Lewy bodies. *Neurology*. 2000; 55(2):302–4. PMID: [10908912](#).
21. Carrieri G, Bonafe M, De Luca M, Rose G, Varcasia O, Bruni A, et al. Mitochondrial DNA haplogroups and APOE4 allele are non-independent variables in sporadic Alzheimer's disease. *Hum Genet*. 2001; 108(3):194–8. PMID: [11354629](#).
22. van der Walt JM, Dementieva YA, Martin ER, Scott WK, Nicodemus KK, Kroner CC, et al. Analysis of European mitochondrial haplogroups with Alzheimer disease risk. *Neurosci Lett*. 2004; 365(1):28–32. <https://doi.org/10.1016/j.neulet.2004.04.051> PMID: [15234467](#).
23. Coskun P, Wyrembak J, Schriener SE, Chen HW, Marciniak C, Laferla F, et al. A mitochondrial etiology of Alzheimer and Parkinson disease. *Biochim Biophys Acta*. 2012; 1820(5):553–64. <https://doi.org/10.1016/j.bbagen.2011.08.008> PMID: [21871538](#); PubMed Central PMCID: PMC3270155.
24. Hudson G, Nalls M, Evans JR, Breen DP, Winder-Rhodes S, Morrison KE, et al. Two-stage association study and meta-analysis of mitochondrial DNA variants in Parkinson disease. *Neurology*. 2013; 80(22):2042–8. <https://doi.org/10.1212/WNL.0b013e318294b434> PMID: [23645593](#); PubMed Central PMCID: PMC3716399.
25. Pyle A, Foltynie T, Tiangyou W, Lambert C, Keers SM, Allcock LM, et al. Mitochondrial DNA haplogroup cluster UKJT reduces the risk of PD. *Ann Neurol*. 2005; 57(4):564–7. <https://doi.org/10.1002/ana.20417> PMID: [15786469](#).
26. Verge B, Alonso Y, Valero J, Miralles C, Vilella E, Martorell L. Mitochondrial DNA (mtDNA) and schizophrenia. *Eur Psychiatry*. 2011; 26(1):45–56. <https://doi.org/10.1016/j.eurpsy.2010.08.008> PMID: [20980130](#).
27. Hagen CM, Aidt FH, Hedley PL, Jensen MK, Havndrup O, Kanters JK, et al. Mitochondrial haplogroups modify the risk of developing hypertrophic cardiomyopathy in a Danish population. *PloS one*. 2013; 8(8):e71904. <https://doi.org/10.1371/journal.pone.0071904> PMID: [23940792](#); PubMed Central PMCID: PMC3734310.
28. Castro MG, Huerta C, Reguero JR, Soto MI, Domenech E, Alvarez V, et al. Mitochondrial DNA haplogroups in Spanish patients with hypertrophic cardiomyopathy. *Int J Cardiol*. 2006; 112(2):202–6. <https://doi.org/10.1016/j.ijcard.2005.09.008> PMID: [16313983](#).
29. Fernandez-Caggiano M, Barallobre-Barreiro J, Rego-Perez I, Crespo-Leiro MG, Paniagua MJ, Grille Z, et al. Mitochondrial haplogroups H and J: risk and protective factors for ischemic cardiomyopathy. *PloS one*. 2012; 7(8):e44128. <https://doi.org/10.1371/journal.pone.0044128> PMID: [22937160](#); PubMed Central PMCID: PMC3429437.

30. Kenney MC, Chwa M, Atilano SR, Falatoonzadeh P, Ramirez C, Malik D, et al. Molecular and bioenergetic differences between cells with African versus European inherited mitochondrial DNA haplogroups: implications for population susceptibility to diseases. *Biochim Biophys Acta*. 2014; 1842(2):208–19. <https://doi.org/10.1016/j.bbadis.2013.10.016> PMID: 24200652; PubMed Central PMCID: PMC4326177.
31. Larsen S, Diez-Sanchez C, Rabol R, Ara I, Dela F, Helge JW. Increased intrinsic mitochondrial function in humans with mitochondrial haplogroup H. *Biochim Biophys Acta*. 2014; 1837(2):226–31. <https://doi.org/10.1016/j.bbabi.2013.10.009> PMID: 24184346.
32. Atilano SR, Malik D, Chwa M, Caceres-Del-Carpio J, Nesburn AB, Boyer DS, et al. Mitochondrial DNA variants can mediate methylation status of inflammation, angiogenesis and signaling genes. *Hum Mol Genet*. 2015; 24(16):4491–503. <https://doi.org/10.1093/hmg/ddv173> PMID: 25964427; PubMed Central PMCID: PMC4512622.
33. Kenney MC, Chwa M, Atilano SR, Falatoonzadeh P, Ramirez C, Malik D, et al. Inherited mitochondrial DNA variants can affect complement, inflammation and apoptosis pathways: insights into mitochondrial-nuclear interactions. *Hum Mol Genet*. 2014; 23(13):3537–51. <https://doi.org/10.1093/hmg/ddu065> PMID: 24584571; PubMed Central PMCID: PMC4049308.
34. Manco J. *Ancestral Journeys*. London: Thames & Hudson; 2015.
35. Cunliffe B. *Europe between the oceans*. New Haven and London: Yale University Press; 2008.
36. Bellwood P. *The Global prehistory of Human Migration*. West Sussex, UK: Wiley Blackwell; 2013.
37. Leslie S, Winney B, Hellenthal G, Davison D, Boumertit A, Day T, et al. The fine-scale genetic structure of the British population. *Nature*. 2015; 519(7543):309–14. <https://doi.org/10.1038/nature14230> PMID: 25788095; PubMed Central PMCID: PMC4632200.
38. Norgaard-Pedersen B, Hougaard DM. Storage policies and use of the Danish Newborn Screening Biobank. *J Inherit Metab Dis*. 2007; 30(4):530–6. <https://doi.org/10.1007/s10545-007-0631-x> PMID: 17632694.
39. Kivisild T. Maternal ancestry and population history from whole mitochondrial genomes. *Investig Genet*. 2015; 6:3. <https://doi.org/10.1186/s13323-015-0022-2> PMID: 25798216; PubMed Central PMCID: PMC4367903.
40. Campbell MC, Tishkoff SA. The evolution of human genetic and phenotypic variation in Africa. *Curr Biol*. 2010; 20(4):R166–73. <https://doi.org/10.1016/j.cub.2009.11.050> PMID: 20178763; PubMed Central PMCID: PMC2945812.
41. Campbell MC, Hirbo JB, Townsend JP, Tishkoff SA. The peopling of the African continent and the diaspora into the new world. *Curr Opin Genet Dev*. 2014; 29:120–32. <https://doi.org/10.1016/j.gde.2014.09.003> PMID: 25461616; PubMed Central PMCID: PMC4308437.
42. Pihl K, Larsen T, Jonsson L, Hougaard D, Krebs L, Norgaard-Pedersen B, et al. [Quality control of prenatal screening]. *Ugeskrift for læger*. 2008; 170(35):2691–5. PMID: 18761859.
43. Pinhasi R, Thomas MG, Hofreiter M, Currat M, Burger J. The genetic history of Europeans. *Trends Genet*. 2012; 28(10):496–505. <https://doi.org/10.1016/j.tig.2012.06.006> PMID: 22889475.
44. Soares P, Achilli A, Semino O, Davies W, Macaulay V, Bandelt HJ, et al. The archaeogenetics of Europe. *Curr Biol*. 2010; 20(4):R174–83. <https://doi.org/10.1016/j.cub.2009.11.054> PMID: 20178764.
45. Benn M, Schwartz M, Nordestgaard BG, Tybjaerg-Hansen A. Mitochondrial haplogroups: ischemic cardiovascular disease, other diseases, mortality, and longevity in the general population. *Circulation*. 2008; 117(19):2492–501. <https://doi.org/10.1161/CIRCULATIONAHA.107.756809> PMID: 18458168.
46. Mikkelsen M, Sorensen E, Rasmussen EM, Morling N. Mitochondrial DNA HV1 and HV2 variation in Danes. *Forensic Sci Int Genet*. 2010; 4(4):e87–8. <https://doi.org/10.1016/j.fsigen.2009.07.007> PMID: 20457038.
47. Li S, Besenbacher S, Li Y, Kristiansen K, Grarup N, Albrechtsen A, et al. Variation and association to diabetes in 2000 full mtDNA sequences mined from an exome study in a Danish population. *Eur J Hum Genet*. 2014; 22(8):1040–5. Epub 2014/01/23. <https://doi.org/10.1038/ejhg.2013.282> PMID: 24448545; PubMed Central PMCID: PMC4350597.
48. Fu Q, Mittnik A, Johnson PLF, Bos K, Lari M, Bollongino R, et al. A revised timescale for human evolution based on ancient mitochondrial genomes. *Curr Biol*. 2013; 23(7):553–9. <https://doi.org/10.1016/j.cub.2013.02.044> PMID: 23523248; PubMed Central PMCID: PMC45036973.
49. Saillard J, Forster P, Lynnerup N, Bandelt HJ, Norby S. mtDNA variation among Greenland Eskimos: the edge of the Beringian expansion. *Am J Hum Genet*. 2000; 67(3):718–26. <https://doi.org/10.1086/303038> PMID: 10924403; PubMed Central PMCID: PMC451287530.
50. Helgason A, Palsson G, Pedersen HS, Angulalik E, Gunnarsdottir ED, Yngvadottir B, et al. mtDNA variation in Inuit populations of Greenland and Canada: migration history and population structure. *Am J Phys Anthropol*. 2006; 130(1):123–34. <https://doi.org/10.1002/ajpa.20313> PMID: 16353217.

51. Starikovskaya YB, Sukernik RI, Schurr TG, Kogelnik AM, Wallace DC. mtDNA diversity in Chukchi and Siberian Eskimos: implications for the genetic history of Ancient Beringia and the peopling of the New World. *Am J Hum Genet.* 1998; 63(5):1473–91. <https://doi.org/10.1086/302087> PMID: 9792876; PubMed Central PMCID: PMCPMC1377559.
52. Raghavan M, DeGiorgio M, Albrechtsen A, Moltke I, Skoglund P, Korneliussen TS, et al. The genetic prehistory of the New World Arctic. *Science.* 2014; 345(6200):1255832. <https://doi.org/10.1126/science.1255832> PMID: 25170159.
53. Underhill PA, Kivisild T. Use of y chromosome and mitochondrial DNA population structure in tracing human migrations. *Annu Rev Genet.* 2007; 41:539–64. <https://doi.org/10.1146/annurev.genet.41.110306.130407> PMID: 18076332.
54. Pakendorf B, Stoneking M. Mitochondrial DNA and human evolution. *Annu Rev Genomics Hum Genet.* 2005; 6:165–83. <https://doi.org/10.1146/annurev.genom.6.080604.162249> PMID: 16124858.
55. Torroni A, Schurr TG, Cabell MF, Brown MD, Neel JV, Larsen M, et al. Asian affinities and continental radiation of the four founding Native American mtDNAs. *Am J Hum Genet.* 1993; 53(3):563–90. PMID: 7688932; PubMed Central PMCID: PMCPMC1682412.
56. Chen YS, Torroni A, Excoffier L, Santachiara-Benerecetti AS, Wallace DC. Analysis of mtDNA variation in African populations reveals the most ancient of all human continent-specific haplogroups. *Am J Hum Genet.* 1995; 57(1):133–49. PMID: 7611282; PubMed Central PMCID: PMCPMC1801234.
57. Torroni A, Huoponen K, Francalacci P, Petrozzi M, Morelli L, Scozzari R, et al. Classification of European mtDNAs from an analysis of three European populations. *Genetics.* 1996; 144(4):1835–50. PMID: 8978068; PubMed Central PMCID: PMCPMC1207732.
58. Emery LS, Magnaye KM, Bigham AW, Akey JM, Bamshad MJ. Estimates of continental ancestry vary widely among individuals with the same mtDNA haplogroup. *Am J Hum Genet.* 2015; 96(2):183–93. <https://doi.org/10.1016/j.ajhg.2014.12.015> PMID: 25620206; PubMed Central PMCID: PMCPMC4320259.
59. Orlando L, Gilbert MT, Willerslev E. Reconstructing ancient genomes and epigenomes. *Nat Rev Genet.* 2015; 16(7):395–408. <https://doi.org/10.1038/nrg3935> PMID: 26055157.
60. Morozova I, Flegontov P, Mikheyev AS, Bruskin S, Asgharian H, Ponomarenko P, et al. Toward high-resolution population genomics using archaeological samples. *DNA Res.* 2016; 23(4):295–310. <https://doi.org/10.1093/dnares/dsw029> PMID: 27436340; PubMed Central PMCID: PMCPMC4991838.
61. Allentoft ME, Sikora M, Sjogren KG, Rasmussen S, Rasmussen M, Stenderup J, et al. Population genomics of Bronze Age Eurasia. *Nature.* 2015; 522(7555):167–72. <https://doi.org/10.1038/nature14507> PMID: 26062507.
62. Fu Q, Posth C, Hajdinjak M, Petr M, Mallick S, Fernandes D, et al. The genetic history of Ice Age Europe. *Nature.* 2016; 534(7606):200–5. <https://doi.org/10.1038/nature17993> PMID: 27135931; PubMed Central PMCID: PMCPMC4943878.
63. Haak W, Lazaridis I, Patterson N, Rohland N, Mallick S, Llamas B, et al. Massive migration from the steppe was a source for Indo-European languages in Europe. *Nature.* 2015; 522(7555):207–11. <https://doi.org/10.1038/nature14317> PMID: 25731166; PubMed Central PMCID: PMCPMC5048219.
64. Lazaridis I, Nadel D, Rollefson G, Merrett DC, Rohland N, Mallick S, et al. Genomic insights into the origin of farming in the ancient Near East. *Nature.* 2016; 536(7617):419–24. <https://doi.org/10.1038/nature19310> PMID: 27459054; PubMed Central PMCID: PMCPMC5003663.
65. Skoglund P, Malmstrom H, Raghavan M, Stora J, Hall P, Willerslev E, et al. Origins and genetic legacy of Neolithic farmers and hunter-gatherers in Europe. *Science.* 2012; 336(6080):466–9. <https://doi.org/10.1126/science.1216304> PMID: 22539720.
66. Larsen NK, Knudsen KL, Krohn CF, Kronborg C, Murray AS, Nielsen OB. Late Quaternary ice sheet, lake and sea history of southwest Scandinavia—a synthesis. *Boreas.* 2009; 38(4):732–61. <https://doi.org/10.1111/j.1502-3885.2009.00101.x> PubMed PMID: WOS:000270642100008.
67. Egeland CP, Nielsen TK, Byo M, Kjaergaard PC, Larsen NK, Riede F. The taphonomy of fallow deer (*Dama dama*) skeletons from Denmark and its bearing on the pre-Weichselian occupation of northern Europe by humans. *Archaeol Anthropol Sci.* 2014; 6(1):31–61. <https://doi.org/10.1007/s12520-013-0117-8> PubMed PMID: WOS:000331997800003.
68. Clark PU, Dyke AS, Shakun JD, Carlson AE, Clark J, Wohlfarth B, et al. The Last Glacial Maximum. *Science.* 2009; 325(5941):710–4. <https://doi.org/10.1126/science.1172873> PMID: 19661421.
69. Price T. D. *Ancient Scandinavia: An Archaeological History From The First Humans To The Vikings.* New York: Oxford University Press; 2015. 1–500 p.

70. Andersson M, Karsten P., Knarrestrom B., Svensson M. Stone Age Scania. Significant Places Dug and Read by Contract Archaeology. Lund: National Heritage Board. Archaeological Excavations, UV Syd.; 2004.
71. Derry TK. History of Scandinavia: Norway, Sweden, Denmark, Finland, and Iceland. Minneapolis, US: University of Minneapolis Press; 2000.
72. Randsborg K. The Anatomy of Denmark. Archaeology and History from the Ice Age to the Present. London: Bristol Classical Press; 2009.
73. Torroni A, Bandelt HJ, D'Urbano L, Lahermo P, Moral P, Sellitto D, et al. mtDNA analysis reveals a major late Paleolithic population expansion from southwestern to northeastern Europe. *Am J Hum Genet.* 1998; 62(5):1137–52. <https://doi.org/10.1086/301822> PMID: 9545392; PubMed Central PMCID: PMCPMC1377079.
74. Petersen VP, Johansen L. Tracking Late Glacial reindeer hunters in Eastern Denmark. *Nationalmuseets Arbejdsmark.* 1996:80–97.
75. Holm J. Settlements of the Hamburgian and Feddermesser culture at Slotseng, South Jylland. *J Dan Arch.* 1993; 10:7–19.
76. Mellars P. A new radiocarbon revolution and the dispersal of modern humans in Eurasia. *Nature.* 2006; 439(7079):931–5. <https://doi.org/10.1038/nature04521> PMID: 16495989.
77. Higham T, Compton T, Stringer C, Jacobi R, Shapiro B, Trinkaus E, et al. The earliest evidence for anatomically modern humans in northwestern Europe. *Nature.* 2011; 479(7374):521–4. <https://doi.org/10.1038/nature10484> PMID: 22048314.
78. Benazzi S, Douka K, Fornai C, Bauer CC, Kullmer O, Svoboda J, et al. Early dispersal of modern humans in Europe and implications for Neanderthal behaviour. *Nature.* 2011; 479(7374):525–8. <https://doi.org/10.1038/nature10617> PMID: 22048311.
79. Posth C, Renaud G, Mitnik A, Drucker DG, Rougier H, Cupillard C, et al. Pleistocene Mitochondrial Genomes Suggest a Single Major Dispersal of Non-Africans and a Late Glacial Population Turnover in Europe. *Curr Biol.* 2016; 26(6):827–33. <https://doi.org/10.1016/j.cub.2016.01.037> PMID: 26853362.
80. Hofmanova Z, Kreutzer S, Hellenthal G, Sell C, Diekmann Y, Diez-Del-Molino D, et al. Early farmers from across Europe directly descended from Neolithic Aegeans. *Proc Natl Acad Sci U S A.* 2016; 113(25):6886–91. <https://doi.org/10.1073/pnas.1523951113> PMID: 27274049; PubMed Central PMCID: PMCPMC4922144.
81. Lazaridis I, Patterson N, Mitnik A, Renaud G, Mallick S, Kirsanow K, et al. Ancient human genomes suggest three ancestral populations for present-day Europeans. *Nature.* 2014; 513(7518):409–13. <https://doi.org/10.1038/nature13673> PMID: 25230663; PubMed Central PMCID: PMCPMC4170574.
82. Melchior L, Lynnerup N, Siegmund HR, Kivisild T, Dissing J. Genetic diversity among ancient Nordic populations. *PloS one.* 2010; 5(7):e11898. <https://doi.org/10.1371/journal.pone.0011898> PMID: 20689597; PubMed Central PMCID: PMCPMC2912848.
83. Fortson B, W. Indo-European Language and Culture: An Introduction. 2 ed. London: Wiley-Blackwell; 2009.
84. Wickham C. The inheritance of Rome: A history of Europe from 400 to 1000. London: Penguin; 2010.
85. Busby GB, Hellenthal G, Montinaro F, Tofanelli S, Bulayeva K, Rudan I, et al. The Role of Recent Admixture in Forming the Contemporary West Eurasian Genomic Landscape. *Curr Biol.* 2015; 25(19):2518–26. <https://doi.org/10.1016/j.cub.2015.08.007> PMID: 26387712; PubMed Central PMCID: PMCPMC4714572.
86. Hvidt K. Danes go west. A book about the emigrations to America. Copenhagen: Rebild National Park Society.; 1976.
87. Brotherton P, Haak W, Templeton J, Brandt G, Soubrier J, Jane Adler C, et al. Neolithic mitochondrial haplogroup H genomes and the genetic origins of Europeans. *Nat Commun.* 2013; 4:1764. <https://doi.org/10.1038/ncomms2656> PMID: 23612305; PubMed Central PMCID: PMCPMC3978205.
88. Malyarchuk B, Derenko M, Grzybowski T, Perkova M, Rogalla U, Vanecsek T, et al. The peopling of Europe from the mitochondrial haplogroup U5 perspective. *PloS one.* 2010; 5(4):e10285. <https://doi.org/10.1371/journal.pone.0010285> PMID: 20422015; PubMed Central PMCID: PMCPMC2858207.
89. Reidla M, Kivisild T, Metspalu E, Kaldma K, Tambets K, Tolk HV, et al. Origin and diffusion of mtDNA haplogroup X. *Am J Hum Genet.* 2003; 73(5):1178–90. <https://doi.org/10.1086/379380> PMID: 14574647; PubMed Central PMCID: PMCPMC1180497.
90. Comas D, Plaza S, Wells RS, Yuldaseva N, Lao O, Calafell F, et al. Admixture, migrations, and dispersals in Central Asia: evidence from maternal DNA lineages. *Eur J Hum Genet.* 2004; 12(6):495–504. <https://doi.org/10.1038/sj.ejhg.5201160> PMID: 14872198.
91. Metspalu M, Kivisild T, Metspalu E, Parik J, Hudjashov G, Kaldma K, et al. Most of the extant mtDNA boundaries in south and southwest Asia were likely shaped during the initial settlement of Eurasia by

- anatomically modern humans. *BMC Genet.* 2004; 5:26. <https://doi.org/10.1186/1471-2156-5-26> PMID: 15339343; PubMed Central PMCID: PMC516768.
92. Gonder MK, Mortensen HM, Reed FA, de Sousa A, Tishkoff SA. Whole-mtDNA genome sequence analysis of ancient African lineages. *Mol Biol Evol.* 2007; 24(3):757–68. <https://doi.org/10.1093/molbev/msl209> PMID: 17194802.
93. Tishkoff SA, Gonder MK, Henn BM, Mortensen H, Knight A, Gignoux C, et al. History of click-speaking populations of Africa inferred from mtDNA and Y chromosome genetic variation. *Mol Biol Evol.* 2007; 24(10):2180–95. <https://doi.org/10.1093/molbev/msm155> PMID: 17656633.
94. Cerezo M, Achilli A, Olivieri A, Perego UA, Gomez-Carballa A, Brisighelli F, et al. Reconstructing ancient mitochondrial DNA links between Africa and Europe. *Genome Res.* 2012; 22(5):821–6. <https://doi.org/10.1101/gr.134452.111> PMID: 22454235; PubMed Central PMCID: PMC3337428.
95. Zupan A, Hauptman N, Glavac D. The maternal perspective for five Slovenian regions: The importance of regional sampling. *Ann Hum Biol.* 2016; 43(1):57–66. <https://doi.org/10.3109/03014460.2015.1006678> PMID: 26065896.
96. Athanasiadis G, Cheng JY, Vilhjalmsen BJ, Jorgensen FG, Als TD, Le Hellard S, et al. Nationwide Genomic Study in Denmark Reveals Remarkable Population Homogeneity. *Genetics.* 2016; 204(2):711–22. <https://doi.org/10.1534/genetics.116.189241> PMID: 27535931; PubMed Central PMCID: PMC5068857.
97. Wallace DC. A mitochondrial paradigm of metabolic and degenerative diseases, aging, and cancer: a dawn for evolutionary medicine. *Annu Rev Genet.* 2005; 39:359–407. <https://doi.org/10.1146/annurev.genet.39.110304.095751> PMID: 16285865; PubMed Central PMCID: PMC2821041.
98. Durieux J, Wolff S, Dillin A. The cell-non-autonomous nature of electron transport chain-mediated longevity. *Cell.* 2011; 144(1):79–91. <https://doi.org/10.1016/j.cell.2010.12.016> PMID: 21215371; PubMed Central PMCID: PMC3062502.
99. Castro MG, Terrados N, Reguero JR, Alvarez V, Coto E. Mitochondrial haplogroup T is negatively associated with the status of elite endurance athlete. *Mitochondrion.* 2007; 7(5):354–7. <https://doi.org/10.1016/j.mito.2007.06.002> PMID: 17660050.
100. Ji Y, Zhang AM, Jia X, Zhang YP, Xiao X, Li S, et al. Mitochondrial DNA haplogroups M7b1'2 and M8a affect clinical expression of leber hereditary optic neuropathy in Chinese families with the m.11778G→A mutation. *Am J Hum Genet.* 2008; 83(6):760–8. <https://doi.org/10.1016/j.ajhg.2008.11.002> PMID: 19026397; PubMed Central PMCID: PMC2668067.
101. Samuels DC, Carothers AD, Horton R, Chinnery PF. The power to detect disease associations with mitochondrial DNA haplogroups. *Am J Hum Genet.* 2006; 78(4):713–20. <https://doi.org/10.1086/502682> PMID: 16532401; PubMed Central PMCID: PMC1424681.
102. Raule N, Sevin F, Santoro A, Altilli S, Franceschi C. Association studies on human mitochondrial DNA: methodological aspects and results in the most common age-related diseases. *Mitochondrion.* 2007; 7(1–2):29–38. <https://doi.org/10.1016/j.mito.2006.11.013> PMID: 17306632.
103. Biffi A, Anderson CD, Nalls MA, Rahman R, Sonni A, Cortellini L, et al. Principal-component analysis for assessment of population stratification in mitochondrial medical genetics. *Am J Hum Genet.* 2010; 86(6):904–17. <https://doi.org/10.1016/j.ajhg.2010.05.005> PMID: 20537299; PubMed Central PMCID: PMC3032075.
104. Achilli A, Rengo C, Magri C, Battaglia V, Olivieri A, Scozzari R, et al. The molecular dissection of mtDNA haplogroup H confirms that the Franco-Cantabrian glacial refuge was a major source for the European gene pool. *Am J Hum Genet.* 2004; 75(5):910–8. <https://doi.org/10.1086/425590> PMID: 15382008; PubMed Central PMCID: PMC1182122.
105. Hudson G, Gomez-Duran A, Wilson IJ, Chinnery PF. Recent mitochondrial DNA mutations increase the risk of developing common late-onset human diseases. *PLoS Genet.* 2014; 10(5):e1004369. <https://doi.org/10.1371/journal.pgen.1004369> PMID: 24852434; PubMed Central PMCID: PMC4031051.
106. Catelli ML, Alvarez-Iglesias V, Gomez-Carballa A, Mosquera-Miguel A, Romanini C, Borosky A, et al. The impact of modern migrations on present-day multi-ethnic Argentina as recorded on the mitochondrial DNA genome. *BMC Genet.* 2011; 12:77. <https://doi.org/10.1186/1471-2156-12-77> PMID: 21878127; PubMed Central PMCID: PMC3176197.
107. Clayton DG, Walker NM, Smyth DJ, Pask R, Cooper JD, Maier LM, et al. Population structure, differential bias and genomic control in a large-scale, case-control association study. *Nat Genet.* 2005; 37(11):1243–6. <https://doi.org/10.1038/ng1653> PMID: 16228001.
108. Marchini J, Cardon LR, Phillips MS, Donnelly P. The effects of human population structure on large genetic association studies. *Nat Genet.* 2004; 36(5):512–7. <https://doi.org/10.1038/ng1337> PMID: 15052271.

109. Wallace DC. Genetics: Mitochondrial DNA in evolution and disease. *Nature*. 2016; 535(7613):498–500. Epub 2016/07/08. <https://doi.org/10.1038/nature18902> PMID: 27383787.
110. Aidt FH, Nielsen SM, Kanters J, Pesta D, Nielsen TT, Norremolle A, et al. Dysfunctional mitochondrial respiration in the striatum of the Huntington's disease transgenic R6/2 mouse model. *PLoS Curr*. 2013; 5. Epub 2013/04/10. <https://doi.org/10.1371/currents.hd.d8917b4862929772c5a2f2a34ef1c201> PMID: 23568011; PubMed Central PMCID: PMC3614423.
111. Hagen CM, Aidt FH, Havndrup O, Hedley PL, Jespersgaard C, Jensen M, et al. MT-CYB mutations in hypertrophic cardiomyopathy. *Mol Genet Genomic Med*. 2013; 1(1):54–65. Epub 2014/02/06. <https://doi.org/10.1002/mgg3.5> PMID: 24498601; PubMed Central PMCID: PMC3614423.
112. Arduino DM, Esteves AR, Swerdlow RH, Cardoso SM. A cybrid cell model for the assessment of the link between mitochondrial deficits and sporadic Parkinson's disease. *Methods Mol Biol*. 2015; 1265:415–24. Epub 2015/01/31. https://doi.org/10.1007/978-1-4939-2288-8_31 PMID: 25634293; PubMed Central PMCID: PMC3614423.
113. Lin YF, Schulz AM, Pellegrino MW, Lu Y, Shaham S, Haynes CM. Maintenance and propagation of a deleterious mitochondrial genome by the mitochondrial unfolded protein response. *Nature*. 2016; 533(7603):416–9. Epub 2016/05/03. <https://doi.org/10.1038/nature17989> PMID: 27135930; PubMed Central PMCID: PMC3614423.
114. Agerbo E, Mortensen PB, Wiuf C, Pedersen MS, McGrath J, Hollegaard MV, et al. Modelling the contribution of family history and variation in single nucleotide polymorphisms to risk of schizophrenia: a Danish national birth cohort-based study. *Schizophr Res*. 2012; 134(2–3):246–52. <https://doi.org/10.1016/j.schres.2011.10.025> PMID: 22108675.
115. Bandelt HJ, Forster P, Rohl A. Median-joining networks for inferring intraspecific phylogenies. *Mol Biol Evol*. 1999; 16(1):37–48. <https://doi.org/10.1093/oxfordjournals.molbev.a026036> PMID: 10331250.
116. Bandelt HJ, Forster P, Sykes BC, Richards MB. Mitochondrial portraits of human populations using median networks. *Genetics*. 1995; 141(2):743–53. PMID: 8647407; PubMed Central PMCID: PMC3614423.
117. Alexander DH, Novembre J, Lange K. Fast model-based estimation of ancestry in unrelated individuals. *Genome Res*. 2009; 19(9):1655–64. <https://doi.org/10.1101/gr.094052.109> PMID: 19648217; PubMed Central PMCID: PMC3614423.
118. Breiman L. Random forests. *Mach Learn*. 2001; 45(1):5–32. <https://doi.org/10.1023/A:1010933404324> PubMed PMID: WOS:000170489900001.
119. Raymond M, Rousset F. An exact test for population differentiation. *Evolution*. 1995; 49(6):1280–3. <https://doi.org/10.1111/j.1558-5646.1995.tb04456.x> PubMed PMID: WOS:A1995TR69500027. PMID: 28568523
120. R: A language and environment for statistical computing. [Internet]. R Foundation for Statistical Computing. 2015. Available from: <http://www.R-project.org>.
121. Patterson N, Price AL, Reich D. Population structure and eigenanalysis. *PLoS Genet*. 2006; 2(12):e190. <https://doi.org/10.1371/journal.pgen.0020190> PMID: 17194218; PubMed Central PMCID: PMC3614423.